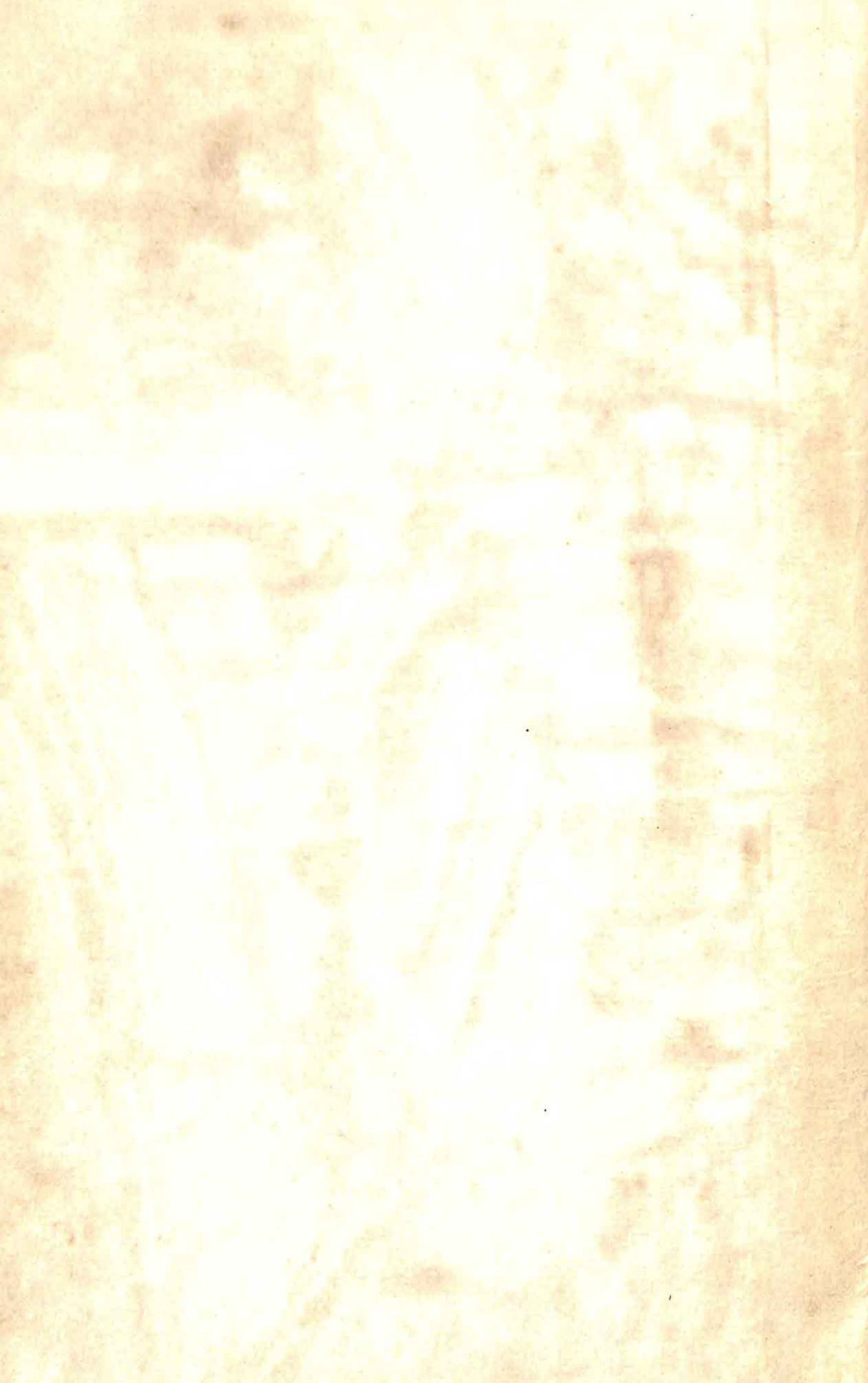


328
3.9.20



(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

EDITED BY

CARL MURCHISON

GORDON ALLPORT
Harvard University

ARTHUR G. BILLS
University of Cincinnati

E. A. BOTT
University of Toronto

LEONARD CARMICHAEL
Smithsonian Institution

HERBERT S. CONRAD
U. S. Office of Education

ELMER K. CULLER
University of Rochester

JOHN F. DASHIELL
University of North
Carolina

ROY M. DORCUS
University of California
at Los Angeles

HORACE B. ENGLISH
Ohio State University

FRANKLIN FEARING
University of California
at Los Angeles

FRANK A. GELDARD
University of Virginia

CLARENCE H. GRAHAM
Columbia University

ERNEST R. HILGARD
Stanford University

HEINRICH KLÜVER
University of Chicago

CARNEY LANDIS
Columbia University

NORMAN L. MUNN
Bowdoin University

JOHN PAUL NAFE
Washington University

DONALD G. PATERSON
University of Minnesota

LEE EDWARD TRAVIS
University of Southern
California

ROBERT C. TRYON
University of California

MICHAEL J. ZIGLER
Wellesley College

If this space should be unstamped, this is the regular library edition.
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not
be resold for a period of five years from date of publication.

VOLUME 49

1960

Copyright, 1960, by The Journal Press
Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.



INDEX OF AUTHORS

Abramson, H. A.	57, 151	Lieberman, Bernhardt	167
Adams, James F.	255	Lifshitz, Kenneth	295
Angers, William P.	123	MacKinnon, William J.	339
Archibald, Herbert C.	155	Maier, Norman R. F.	217
Armengal, Vilma	117	Meade, Robert D.	249
Bartley, S. Howard	3, 185, 263	Meltzer, H.	43
Comalli, Peter E., Jr.	99	Miles, Guy H.	233
Crowder, T. H.	181	Nafe, John Paul	29
Crowder, William F.	83, 181	Nelson, Thomas M.	3, 185, 263
Currier, W. D.	67	Newsome, George L.	333
Curti, Margaret Wooster	13	Nogee, Philip	167
Dawson, W. W.	29	Phillips, E. Lakin	195
DeHardt, Doris	3, 263	Porter, Paul B.	139
Dittborn, Julio M.	111, 113, 117	Robertson, Malcolm H.	349
Francesco, E.	213	Rolo, A.	151
Gilley, Philip F. M., Jr.	289	Ross, Bruce M.	133
Gladin, Leo L.	305	Rust, Ralph M.	235
Godoy, L. M.	111, 113	Sklarofsky, B.	151
Gough, Harrison G.	87	Stache, J.	151
Gutierrez, O.	111, 113	Thompson, Clare W.	155
Hawkes, Glenn R.	203, 279	Tipper, Harry	145
Jencks, Beata	139	Warm, Joel S.	279
Jones, Frank Pierce	289	Watson, George	67
Kenshalo, D. R.	29	Wilkes, William P.	83, 181
Kingston, Albert J.	333	Woodworth, Donald G.	87
Levy, Nissim	133		

3.9.70
J-328

TABLE OF CONTENTS

A comparison of variability of three sorts of observers in a sensory experiment BY THOMAS M. NELSON, S. HOWARD BARTLEY, AND DORIS DEHARDT	3
Intelligence tests of white and colored school children in Grand Cayman BY MARGARET WOOSTER CURTI	13
A new method for the investigation of thermal sensitivity BY D. R. KENSHALO, JOHN PAUL NAFE, AND W. W. DAWSON	29
Scope of industrial psychology references in the <i>Annual Review</i> , 1950-1959 BY H. MELTZER	43
Lysergic Acid Diethylamide (LSD-25): XXX. The questionnaire technique with notes on its use BY H. A. ABRAMSON	57
Intensive vitamin therapy in mental illness BY GEORGE WATSON AND W. D. CURRIER	67
Secondary reinforcement with and without control for response facilitation BY WILLIAM P. WILKES AND WILLIAM F. CROWDER	83
Stylistic variations among professional research scientists BY HARRISON G. GOUGH AND DONALD G. WOODWORTH	87
Studies in physiognomic perception: VI. Differential effects of directional dynamics of pictured objects on real and apparent motion in artists and chemists BY PETER E. COMALLI, JR.	99
Sleep Suggestibility Test BY JULIO M. DITTBORN, O. GUTIERREZ, AND L. M. GODOY	111
Expectation as a factor of sleep suggestibility: II BY JULIO M. DITTBORN AND VILMA ARMENGOL	113
An operational definition of somnambulist hypnosis BY JULIO M. DITTBORN AND VILMA ARMENGOL	117
Job counseling of the epileptic BY WILLIAM P. ANGERS	123
A comparison of adjectival antonyms by simple card-pattern formation BY BRUCE M. ROSS AND NISSIM LEVY	133
Need reduction and primary reinforcement: Varied goal situations as reinforcement in maze learning of rats BY BEATA JENCKS AND PAUL B. PORTER	139
Effect of psychostimuli on technical personnel BY HARRY TIPPER	145
Production of cross-tolerance to psychosis-producing doses of Lysergic Acid Diethylamide and Psilocybin BY H. A. ABRAMSON, A. ROLO, B. SKLAROFSKY, AND J. STACHE	151
The span of human development BY HERBERT C. ARCHIBALD AND CLARE W. THOMPSON	155

The auction value of certain risky situations	167
BY PHILIP NOGEE AND BERNHARDT LIEBERMAN	
Weak-light reinforcement with and without control for response facilitation .	181
BY W. F. CROWDER, W. P. WILKES, AND T. H. CROWDER	
A comparison of three rates of pulse onset and decline in producing critical flicker frequency	185
BY S. HOWARD BARTLEY AND THOMAS M. NELSON	
Parent-child psychotherapy: a follow-up study comparing two techniques .	195
BY E. LAKIN PHILLIPS	
Cutaneous communication: Absolute identification of electrical intensity level	203
BY GLENN R. HAWKES	
The verbal power test of concept equivalents (VPT)	213
BY E. FRANCESCO	
Screening solutions to upgrade quality: A new approach to problem solving under conditions of uncertainty	217
BY NORMAN R. F. MAIER	
Modification of a commonly used source of electric shock	233
BY GUY H. MILES	
Epidemiology of mental health in college	235
BY RALPH M. RUST	
Time perceptions as affected by need tension	249
BY ROBERT D. MEADE	
Test item difficulty and the reliability of item analysis methods	255
BY JAMES F. ADAMS	
Considerations involved in obtaining averages of <i>CFF</i> data	263
BY THOMAS M. NELSON, S. HOWARD BARTLEY, AND DORIS DEHARDT	
Maximum I_t for absolute identification of cutaneous electrical intensity level	279
BY GLENN R. HAWKES AND JOEL S. WARM	
Head-balance and sitting posture: An x-ray analysis	289
BY FRANK PIERCE JONES AND PHILIP F. M. GILLEY, JR.	
Problems in the quantitative evaluation of patients with psychoses of the senium	295
BY KENNETH LIFSHITZ	
The stimulus concept: <i>Animism</i> ?	305
BY LEO L. GLADIN	
The relationship of two measures of authoritarianism to the Minnesota Teacher Attitude Inventory	333
BY ALBERT J. KINGSTON AND GEORGE L. NEWSOME	
Behavioral research and intergroup communication	339
BY WILLIAM J. MACKINNON	
The student's need for an issue	349
BY MALCOLM H. ROBERTSON	

A COMPARISON OF VARIABILITY OF THREE SORTS OF OBSERVERS IN A SENSORY EXPERIMENT*¹

Department of Psychology, Michigan State University

THOMAS M. NELSON, S. HOWARD BARTLEY, AND DORIS DEHARDT

A. PROBLEM

A persistent problem in psychology has to do with what to expect from naïve and trained observers. Although various schools of endeavor have long used trained observers, the belief has not totally disappeared in certain quarters that trained observers do not yield the type of data the investigator supposes he is getting, but that naïve observers do meet the proper qualifications. It is supposed that the training process is unavoidably a biasing process, and that the biasing is in favor of obtaining results that the investigator wants. The interpretation to be made here briefly is that a form of self-deception is being practiced by investigators confining their work to the use of trained observers. One of the authors (Bartley), in his years of teaching psychology, has gained the impression that students in his classes have somewhere been taught to frown on trained observers.

It is known that reports (judgments, observations, or whatever) vary from trial to trial. This is always the case and has led to statistical treatment of the results, all the way from simple averaging of the readings and the use of the arithmetic mean as a single representative reading to more elaborate procedures such as used in the method of constant stimuli. It is also known that variability differs under fairly well known diverse circumstances such as from day to day, or from hour to hour as contrasted with results when readings are taken in rapid succession, as from trial to trial.

The more serious the investigator, the more likely he attempts to consider just what is occurring when human observations are being made. He wishes to ask himself what needs to be done to stabilize his observer's responses, or to do something acceptable regarding variations that seem to be unavoidable.

It has become customary to rely on statistical procedures to determine amounts of variation, to state variation in a standard way, and to tell something about the value of the results obtained. By value is meant the repre-

* Received in the Editorial Office on September 1, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The present study is a part of the work being done under Research Grant NSF-G-5821 from the National Science Foundation.

sentativeness of the results, or the expectancy of being repeatable. Frequently the way statistics is used in psychology implies an evaluation that lies totally within the number of scheme (mathematics) employed, rather than in understandings and/or assumptions aside from the scheme. The resort to statistics has been so common as to bespeak a widespread belief that there are no other criteria for evaluating experimental results. It is as if we now have hit upon a new and much more effective way of dealing with nature than any heretofore in use. On the other hand, it must be recognized that there is yet something further to be achieved in our conceptualization of what we are doing in a sensory experiment. This understanding lies outside of the application of mathematical formalities. It involves more intimate understandings of body mechanisms.

The situation evoking the present inquiry has stemmed from the need of obtaining considerable data on human response to intermittent visual stimulation, i.e., flicker. In this investigation, like many others, one needs to be able to compare (*a*) the behavior of individuals who have had extended experience with (*b*) those who have been briefly though carefully instructed and trained, with (*c*) those whom we could well call naïve. The present paper presents the results of just such a comparison.

B. OBSERVER MATERIAL AND PROCEDURE

In making comparisons, naturally a standard or reference is necessarily involved. In the present case, no absolute or external standard was available. Hence the standard of variability we used was the behavior of the two *O*s whom we label as trained *O*s, meaning that they have had extended experience in sensory observation and in the consideration of the meaning of the endeavors in which they have been engaged. These *O*s were two of the authors.

The second group of *O*s (sophisticates) were six individuals who underwent extended instruction in regard to just what they were to observe and what they were to ignore in the visual situation, as a part of a course in experimental psychology they were taking. These persons made their observations over a 10-week period.

The third group of observers (naïve) was students from a class in elementary psychology who had never before served in a sensory investigation. These persons were, of course, given verbal instruction, but only the kind usually given to obtain readings in an observational experiment for the first time. The total time spent in observations for individuals in this group totaled only two hours.

For all observers, the task consisted of viewing an intermittent target,

whose rate of intermittency was varied by the observer himself until all vestige of flicker just disappeared. Four different intensity levels and seven different *PCF*'s (pulse-to-cycle fractions) were used.

Both of the trained *O*s and the sophisticates observed under all of the seven *PCF* conditions ($1/30$, $1/8$, $1/4$, $1/2$, $3/4$, $7/8$ and $29/30$) as well as under the four intensity levels. No single naïve observer served for more than two *PCF* conditions, though all four of the intensities were used. This reduced the experience that any naïve observer had at the time of his very last trials.

C. RESULTS

The results are presented in two forms, the graphical in Figures 1, 2, and 3, and the statistical. In Figures 1 and 2, one can see the behavior of individuals in the three groups quite distinctly. The points representing the sophisticated *O*s are shown by solid circles and those for the naïve by open circles. The mean responses for the trained *O*s are shown by triangles and crosses. The difference between the sophisticated and the naïve observers

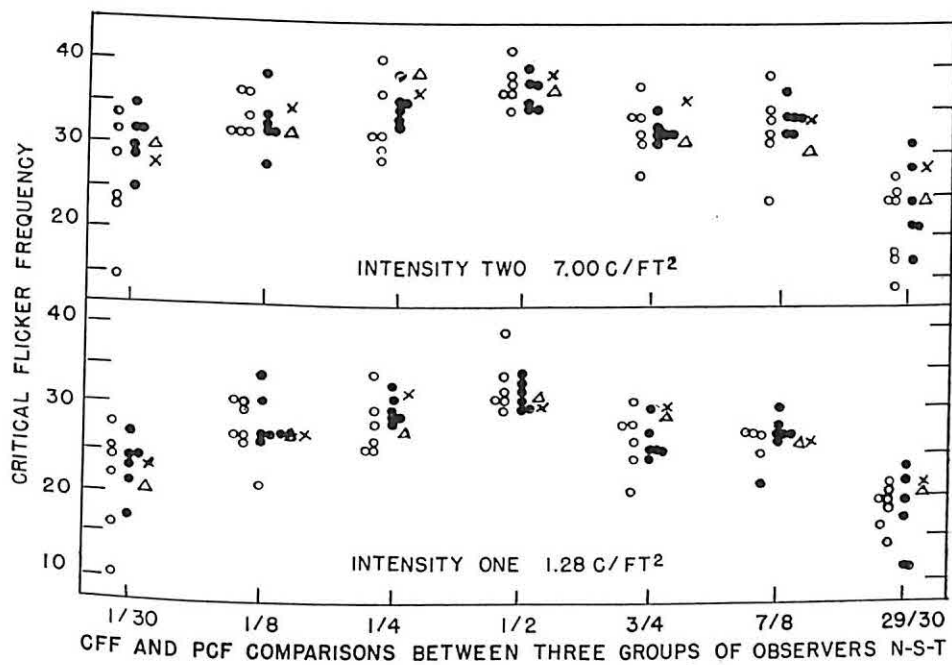


FIGURE 1

Scatter diagram of critical flicker frequency responses of three groups of *O*s for two intensities and seven pulse-to-cycle fractions. Sophisticated *O*s shown in solid circles, naïve *O*s in open circles. Responses of the two trained observers shown by triangles and crosses, respectively.

consisted largely in the amounts of scatter from observer to observer. It will be shown that statistically all three groups differed in amounts of variability. It can also be seen from the scatters of points that, in most cases, the scatter is distributed in both directions from the values represented by the two trained observers. Figure 3 shows the combined results of the individuals in each of the three groups. For each of four different intensity levels, three curves are presented. The solid curves represent the trained *O*s; the finely broken curves, the sophisticated *O*s; and the grossly broken curves, the naïve *O*s.

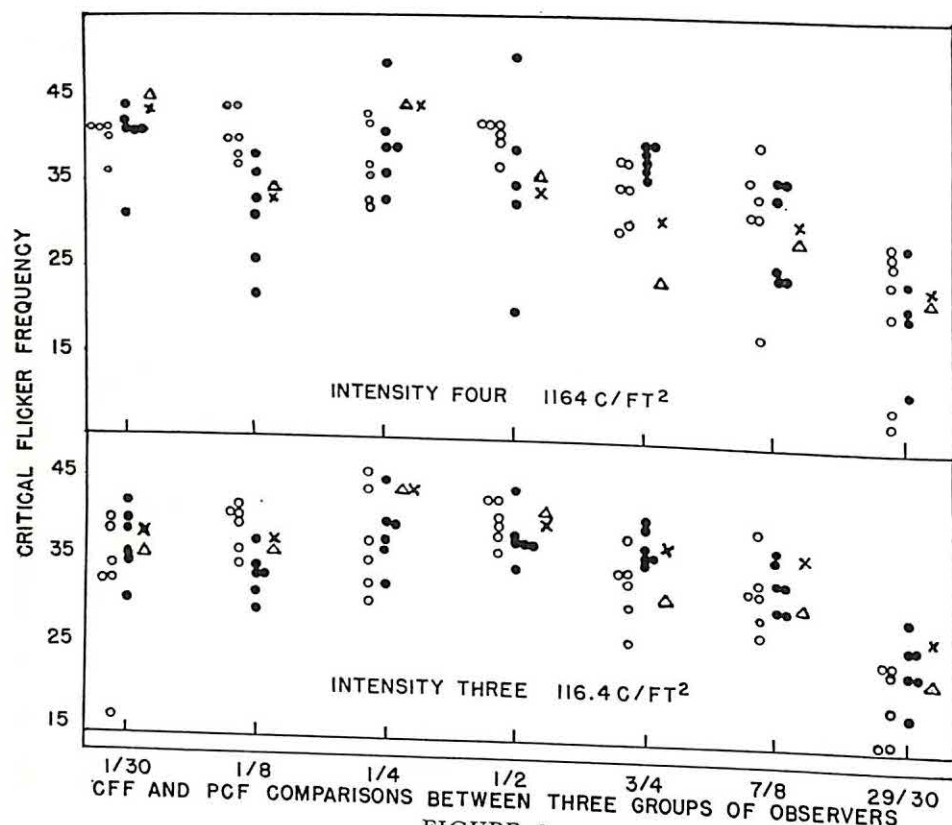


FIGURE 2

Scatter diagram of critical flicker frequency responses of three groups of *O*s for two intensities and seven pulse-to-cycle fractions. Sophisticated *O*s in open circles. Responses of the two trained observers shown by solid circles, and crosses, respectively.

It will be noted that at the lowest intensity level, all three groups behaved quite similarly. A slightly greater degree of divergence occurred at the next higher level. At the third level this divergence increased, and at the highest

level it became marked. One of the very most effective factors that would seem to play a rôle in this divergence is the nature of the target and the stray photic radiation produced in the eye. At the very highest intensity, the field surrounding the target is seen to flicker violently and it requires extreme skill and caution not to be influenced by this. This field flicker obscures the observation of the last vestiges of flicker in the target. If the *O* fails to ignore this flicker, he increases the intermittency rate beyond the values he should, be-

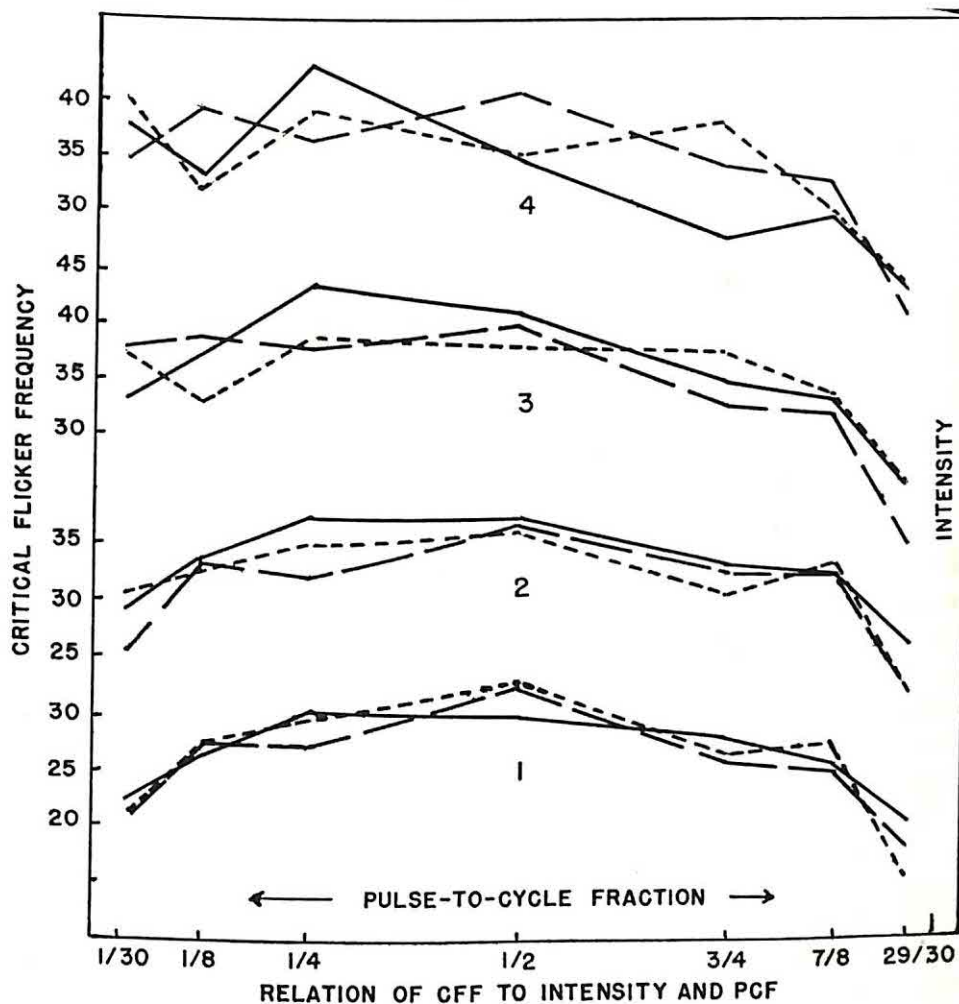


FIGURE 3

CFF curves for three groups of *O*s. Solid lines are for trained observers, finely broken lines for naïve *O*s, and grossly broken lines for sophisticated *O*s. The numbers represent intensity levels. Number 1 = 1.28 c/ft², 2 = 7.00 c/ft², 3 = 1/16.4 cft², and 4 = 1164 c/ft².

cause he is obtaining the *CFF* of the field instead of the target. Some years ago, Bartley (1) studied the relation of *CFF*'s for target and surrounding field and showed that, under some conditions, the field continued to flicker with pulse rates higher than those producing *CFF* for the target. Conditions for field flicker were studied and it was shown that it was produced by stray photic radiation in the eye. This radiation cannot be fully reduced to negligibility under certain conditions.

The statistical treatment of the material was as follows: Pooled or average variances (s_a^2) for each of the three populations of *O*s were found and corresponding *F* tests were performed. For both trained *O*s, 10 scores were obtained for each of seven *PCF*'s at each of four intensity levels. That is, 28 samples consisting of 10 scores each were obtained for both *O*s. Thus, there were 56 trained *O* samples (560 scores). The variance (s^2) of each of these samples was computed. Since all of these samples were considered to have come from the same population, the resulting 56 variances were averaged to give s_a^2 , s_a^2 being an estimate of the variance of the population of trained *O*s. Twenty-eight samples identical to those for the trained *O*s were also obtained for each sophisticated *O*. Since there were six sophisticated *O*s, the total number of sophisticated *O* samples was 168 (1680 scores). Again, variances for each sample were computed and the average variance obtained.

There were also 168 naïve *O* samples from which an average variance was computed. But unlike *O*s in the other two groups, each naïve *O* was tested on only two *PCF*'s for each of the four intensity levels. Before testing of naïve *O*s began, various combinations of two from the seven possible *PCF*'s were listed in such a way that each *PCF* appeared exactly six times. This was done so that the naïve *O*s' s_a^2 would correspond to s_a^2 for trained and sophisticated *O*s. A total of 21 combinations was given to as many naïve *O*s and with eight samples for each, the number of samples totaled 168. Again, since the 21 *O*s arose from the same population, an average variance was computed.

The average variances just described for the three groups will be called overall average variances (s_o^2), and are averages of *all* samples for each group. In addition to those for observer groups, average variances were found for the various intensities (s_d^2) and *PCF*'s (s_p^2) also. Four average variances were computed for intensities and seven for *PCF*'s for each group. For example, an average variance for trained *O*s at Intensity 2 (I_1) would consist of all trained *O* samples at I_1 .

Presented in Table 1 are the overall average variances (s_o^2) and the *F* ratios of these variances. Both *F* ratios are ratios of sophisticated and naïve

s_o^2 to trained s_o^2 . For $F = \text{sophisticated } s_o^2 / \text{trained } s_o^2$, 2.16 ($df = 1512, 504$) was obtained, which is significant at $p < .01$. The ratio of naïve s_o^2 to trained s_o^2 is 2.43 ($df = 1512, 504$) an F ratio also significant at $p < .01$. The trained s_o^2 is very significantly smaller than either the sophisticated s_o^2 or the naïve s_o^2 , both results supporting the hypothesis.

TABLE 1
OVERALL AVERAGE VARIANCES (s_o^2) AND F RATIOS FOR TRAINED, SOPHISTICATED, AND NAÏVE OBSERVERS

Observers	s_o^2	df	F
Trained	111.9	504	
Sophisticated	241.7	1512	2.16*
Naïve	271.5	1512	2.43*

* Significant at $p < .01$.

Not presented in Table 1 but also computed was $F = \text{naïve } s_o^2 / \text{sophisticated } s_o^2 = 1.12$ ($df = 1512, 1512$), which is significant at $p < .05$. The F distribution percentile for $\alpha = .05$ for such a large df was approximated by the technique given by Dixon and Massey (2). In other words, the sophisticated s_o^2 is significantly smaller than the naïve s_o^2 , a result which was also hypothesized.

TABLE 2
AVERAGE VARIANCES FOR DENSITIES (s_d^2) AND F RATIOS FOR TRAINED, SOPHISTICATED, AND NAÏVE OBSERVERS

AND NAÏVE OBSERVERS					
Observers		I ₁	I ₂	I ₃	I ₄
		s _d ²			
Trained		60.4	81.8	111.5	194.1
Sophisticated		93.5	164.8	193.9	514.7
Naïve		351.6	239.1	247.5	247.8
		<i>F Ratios</i>			
Observers	<i>df</i>	F _{I₁}	F _{I₂}	F _{I₃}	F _{I₄}
Trained	126				
Sophisticated	378	1.55*	2.01*	1.74*	2.65*
Naïve	378	5.82*	2.92*	2.22*	1.28

*Significant at $p < .01$.

It can be said, then, that the hypothesis that trained O s are less variable than the sophisticated and the naïve O s, and that the sophisticated O s are less variable than the naïve O s is verified.

The average variances for intensities (s_d^2) are presented in Table 2. F_{I_1} is the F ratio for Intensity 1, etc. Again, the ratios compared naïve and

sophisticated s_d^2 to trained s_d^2 . All trained s_d^2 are significantly *smaller* ($p < .01$) than those of naïve and sophisticated s_d^2 , *except* F_{I_3} for the naïve O_s .

TABLE 3
AVERAGE VARIANCES FOR PCF 's (s_p^2) AND F RATIOS FOR TRAINED, SOPHISTICATED, AND NAÏVE OBSERVERS

Observers	1/30	1/8	1/4	1/2	3/4	7/8	29/30	
s_p^2								
Trained	101.3	248.2	59.7	66.3	115.6	120.8	71.7	
Sophisticated	228.0	259.7	210.6	171.5	157.5	375.2	289.4	
Naïve	366.6	264.6	272.7	293.2	138.7	264.0	301.0	
F Ratios								
Observers	df	$F_{1/30}$	$F_{1/8}$	$F_{1/4}$	$F_{1/2}$	$F_{3/4}$	$F_{7/8}$	$F_{29/30}$
Trained	72							
Sophisticated	216	2.25*	1.05	3.53*	2.59*	1.36	3.11*	4.04*
Naïve	216	3.62*	1.07	4.57*	4.42*	1.20	2.19*	4.20*

* Significant at $p < .01$.

From Table 3, it can be seen that all average variances for PCF 's (s_p^2) for trained O_s are significantly *smaller* ($p < .01$) than those for naïve and trained O_s s_p^2 for five of the seven PCF 's. For PCF 's 1/8 and 3/4, significant differences were not obtained.

D. DISCUSSION

It is to be admitted that there is more than one way of looking at sensory data. One is the statistical, which, as a purely mathematical scheme, offers ways of treating data and indicating how they shall be regarded. The other is an empirical and concrete way in which data are seen as products of the operation of physicalistic mechanisms, be they biological or be they those outside the body. Variation is looked upon by these two views in somewhat different ways, although, in both, variation is unwanted. Variation is difficult to make the kind of sense out of that some of us look for. To say the least, we see it as an indication that the mechanisms at work are not all brought in line with the specific experimental performance expected of the subject or observer.

Considerable could be said in this connection to make the discussion satisfactory. The purpose for mentioning the two quite different views has been merely to indicate that the statistical results obtained in the present study are not taken as a final objective, but rather only as a way of stating a kind of equivalence or lack of it between three groups of observers. The data answer such questions as (a) whether one group might be substitutable for either of

the others, and (b) whether, although amounts of variability differ, all groups are essentially providing the same overall pattern of data. It will be remembered that in the introduction we voiced the doubt, existing in some quarters, whether naïve and uninformed observers and the trained and informed ones address themselves to essentially the same task. The results we have obtained should tend to dispel the idea that phenomena such as flicker and the *CFF* are highly artifactual and arbitrary phenomena.

It would seem that although the three groups differed significantly in amounts of variability, they did demonstrate that, in general, they were pursuing the same task. If such tasks can be quickly assumed by totally naïve observers, it would seem that they are quite "natural."

In a subsequent paper we shall deal with the question of how many readings are needed per observer per single set of conditions in order to be satisfactory from some reasonable, if not from some formal statistical standpoint.

REFERENCES

1. BARTLEY, S. H. The basis of flicker in the visual field surrounding the test-object. *J. Exper. Psychol.*, 1936, **19**, 342-350.
2. DIXON, W. J., & MASSEY, F. J., JR. *Introduction to Statistical Analysis*. New York: McGraw-Hill, 1957.

Department of Psychology
Michigan State University
East Lansing, Michigan

INTELLIGENCE TESTS OF WHITE AND COLORED SCHOOL CHILDREN IN GRAND CAYMAN*

Madison, Wisconsin

MARGARET WOOSTER CURTI¹

A. PROBLEM

Psychologists have increasingly pointed out that the inferiority of Negroes on intelligence tests can be explained as a result of economic and educational inferiority, but attempts to control such factors have not produced clear-cut results. In a recent careful survey Dr. Audrey M. Shuey reports that at several age levels and under a variety of conditions, Negroes regularly score below whites. In a foreword to this book Henry E. Garrett says: "Dr. Shuey concludes that the regularity and consistency of the results strongly imply a racial basis for these differences. I believe that the weight of evidence supports her conclusion" (1).

One of the main reasons for the inconclusiveness of studies of Negro-white intelligence is that usually selection of one kind or another may be presumed to have affected the composition of the groups compared. Another is that not infrequently tests used have included materials and involved procedures within the experience of one group but not the other. Perhaps the most important factor not adequately controlled lies in the conscious or subconscious attitudes toward those tested of the person giving the tests.

The writer in 1940 carried out in Jamaica a study designed to give more adequate consideration to such factors. As part of that investigation a study was made on the island of Grand Cayman, a dependency of Jamaica. Results were reported to the British government authorities but not published. Since at the present time the findings have special relevance, this paper reports on the group intelligence testing in Cayman.

Grand Cayman seemed to be a natural laboratory for racial testing for several reasons: English is the language spoken; the island is small enough (20 miles long) so that all children of a given age in school could be tested; a sizeable part of the population has remained "pure white" while there are large numbers of brown and black people; and the island at the time of the

* Received in the Editorial Office on September 6, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ Study supported by Social Science Research Council and the Education Office, Jamaica.

study was quite isolated and all the inhabitants were living in an unusually simple environment.

As to the last point, there was no regular steamship service, no railroad, moving picture theater, or newspaper. There were very few automobiles, and very few telephones. The roads were generally poor. There were no government schools above the elementary level and no private schools above the secondary level. The great majority of the people lived off the sea, in one way or another. In Grand Cayman there were many white families which had been exposed for some generations to the same simple type of environment as colored and black families. People of all colors, then, were restricted to a rather narrow range of occupational opportunity.

B. METHOD

1. Preliminary Testing

Some preliminary tests were given under the writer's direction by Mr. Frank Ogle of the Jamaica Education Office during an inspection trip to Cayman. He administered group intelligence tests to government school children, and the results were used as a basis for planning later testing. The only test regularly given to the children present on the day of inspection was the Goodenough draw-a-man test, the results of which are summarized in Table 1.

TABLE 1
RESULTS, PRELIMINARY MAN TEST

Age	N	White		Brown			Black			White-Brown		White-Black	
		M	SD	N	M	SD	N	M	SD	Diff.	t	Diff.	t
8	5	11.66	6.8	10	15.8	8.	5	13.	4.8	-5.2	1.2	-2.4	.57
9	8	16.	6.6	8	14.6	5.9	8	14.	5.7	+1.38	.41	+2.	.6
10	9	21.2	7.0	12	16.8	6.5	8	20.2	5.0	+4.4	1.22	+ .97	.31
11	11	14.6	10.3	15	18.1	6.6	5	19.2	4.1	-3.4	.92	-4.57	1.17
12	14	18.71	9.5	25	18.72		10	19.4	9.1	-.01		-.69	.17

Explanation: A minus sign signifies white mean is lower, a plus sign, higher. No color *t*'s are significant.

These results were striking in that the commonly found inferiority of black pupils was not shown. Black averages are higher than white in 3 out of 5 comparisons. But neither these nor other differences are significant even near the 5 per cent level. Since all the children tested were attending government schools, the writer thought that if private schools could be included, the averages of the white pupils would at least equal and probably exceed those of the black and the colored.

2. *The Main Cayman Study*

In the main study the writer visited Grand Cayman with an assistant and conducted tests in all the schools on the island. The heads of all the private schools gave generous and intelligent help. Both teachers and pupils in all the schools were interested and gave excellent cooperation. Nearly all the actual testing was done by an assistant, a native Jamaican, a young woman with a name common in Cayman. The writer supervised the obtaining of supplementary information.

In all the schools in Grand Cayman we gave group tests to all in the lower standards (grades), checking on school lists to make sure of getting all children through 10 years old according to birth certificates. Whenever time permitted the older pupils (called seniors by us) were also tested, and senior testing was done in all the private and most of the government schools. We excluded from the main treatment the few whose birthdates could not be verified in government records, and also the small number of children aged 6 and 7 years. Table 2 shows the number of pupils dealt with according to color and type of school.

TABLE 2
SUBJECTS

Age	White		Brown		Black		All	
	Private	Govern- ment	Private	Govern- ment	Private	Govern- ment	Private	Govern- ment
Jr. (8-10)	31	28	9	53	7	44	47	125
Sr. (11-15)	27	36	7	66	2	33	36	135
Both	58	64	16	119	9	77	83	260

Whether the 122 white, 135 brown, and 86 black children reflect the racial make-up of the island we can not be sure since census reports have not made a clear distinction between white and very light brown, and since even the darkest are probably less than pure black. Cayman advisers told us they felt this proportion of black and brown children seemed about right.

In regard to the question whether the very dull or defective were in school, it is significant that illiteracy, backwardness, and feeble-mindedness were reported by the teachers. Of 171 juniors, 13 per cent were reported illiterate (15 per cent of the white, 10 per cent of the brown, and 17 per cent of the black pupils). Illiteracy was not usually reported for the older pupils, but one white boy age 11 was said to be illiterate, and such comments as "can read but little" occurred. Among the seniors three were reported feeble-minded and one mentally defective. Apparently not a few pupils unable to do the school work as laid down in the course of study, were kept on any-

way and kept busy in some way. Our figures suggest that Cayman school children were more representative and "unselected" than is common in less simple societies.

Before a school was visited we sent to the teacher a letter suggesting that she interest the children in our visit (tests were not to be mentioned) and ask them to come on the appointed day with birth certificates. Teachers reported great interest and unusually large attendance on the day of testing. In Cayman our activities provided a welcome change.

Supplementary information was obtained for each pupil as follows: height in stocking or bare feet, measured at the time; date and place of birth, legitimacy and also full names of both parents to permit checking by government records; occupation of father and of mother; occupation child would prefer when grown up; number of people living in child's home; ability to read and to write (reported by teacher).

In addition the teacher, in conference with the writer, rated each child, using mimeographed instructions, on these items: schoolwork, intelligence, self-confidence, type of home, and health. The teacher also recorded whether eyesight and hearing were good or not, and was also asked to report any special defect or handicap he had noticed. These last questions elicited mention of various defects, "poor teeth" being the most common, and others occurring such as "bad blood," "nervous twitching," "kidney trouble," etc.

A rating on color was given each pupil on a 6-point scale ranging from 1 (black) through 6 ("pure" white). The writer and her assistant conferred about doubtful cases. An occupational-economic rating was also given, based mainly on reported occupations of both parents but modified in accordance with certain rules by consideration of the teacher's report on the home. Ratings were on a 6-point scale from 1 (highest) to 6 (destitute).

3. *The Tests*

All tests were specially devised or adapted in Jamaica in consultation with native Jamaicans. In addition to the group tests, some 5-year-olds not in school were given special "kindergarten" tests, and all pupils within one month of the ninth birthday were given a modified Binet-type test. In certain schools, time permitting, pupils were asked to "make a picture—anything you want to draw."

In this paper only the group tests are reported on. Material for each test was on a separate mimeographed sheet. Brief descriptions of the tests and reactions to them follow:

a. *Junior Tests.* (1.). For the *draw-a-man* test the standard directions for giving and scoring were used. Mental ages were not calculated.

(2). *Counting blocks.* There were 25 piles of blocks of increasing complexity. Below each pile the pupil was to make one mark for each block in the pile, or if he wanted to, to write below each the number of blocks. Score was number correct. This test, with additional more complex piles, was used for seniors.

(3). The *lost purse test* was an adaptation of the ball and field test. Instead of a pass or fail criterion, scores from 0 to 10 were assigned according to goodness of the planning.

(4). *Pictorial classification.* For each of 25 rows of geometrical figures the pupil was to put a mark on the figure (picture) that was most different from the rest. The first two rows were used for demonstration and practice. With the addition of two more complex rows, the test was used for seniors. Score was number of rows correct.

(5). For the *memory squares* test a row of squares was drawn on the blackboard, and each child had a sheet with 10 rows of blank squares. The tester explained that he would tap certain squares with his ruler, and that when he lowered it, each pupil was to put marks in the same squares on his paper. There was demonstration with two practice rows. The test was received with downright enthusiasm. It was used for seniors but with more complex patterns. Score was number of rows correct.

(6). *Geometrical forms.* In each one of 14 compartments on a sheet, two or more parts of a geometrical form were shown, followed by an outline of the entire figure. The pupil was to show by marking the larger figure how the parts could be arranged to make the figure. Demonstration and explanation by sample blackboard figures were given. Score was number of forms correctly done, with half credit where part but not all of the parts were shown in correct position. The test was also given to seniors.

b. *Senior Tests.* Senior tests 1, 2, 3, and 4 (blocks, pictorial classification, memory squares, and forms) have been described under Junior Tests.

(5). *Opposites.* Two vertical columns of opposites, 50 pairs in all, were presented. After each pair, in space provided, the subject was directed to write an "S" if the two words meant the same or nearly the same, a "D" if they were very different. Score was number right minus number wrong. Pairs were of increasing difficulty, from "1. white-black," to "49. avarice-cupidity," and "50. quite-entirely."

(6). *Verbal classification.* For each of 25 rows of words, the pupil was directed to "Draw a line under the word that is most different from the

rest of the words in the row." Example: "8. inch, century, yard, mile, foot." Score was number of rows correctly marked.

(7). *Number completion*. In each of 25 rows was a series of numbers arranged according to a definite rule. The subject was to write, in the space provided, the two numbers needed to complete each row. Score was number of rows correctly marked. Sample lines: "1, 5, 9, 13, 17, 21, —, —;" "3, 6, 8, 16, 18, 36, —, —."

(8). A conventional *analogies* test was given, the correct word to be selected by underlining one of four different choices. Example: Water is to drink as bread is to (cake, coffee, eat, bun). Score was number of lines correctly marked out of a total of 25.

(9). A test we thought of as the "*reasoning*" test was really two: (a) In each of five statements, one part was left to be completed by the pupil who was to "draw a line under the number or words which best complete the meaning." Example: May is taller than Alice, and Alice is taller than Pearl. So May is [taller than, shorter than, just as tall as] Pearl. (b) Three proverbs common in the islands were presented, each followed by five statements. Directions were, "In the square at the right, put the number of the statement which best explains the proverb." Score on this test was the number of correct responses on both (a) and (b).

C. THE RESULTS

1. *Treatment*

Analysis of variance was the basic method used. Age was one criterion of classification. Among seniors there were not enough pupils aged 14 for all sub-groups, and so 15-year-olds were included too in a "14+" group. Three color-classes were distinguished. The black group included all rated 1 and all rated 2 who *also* had kinky hair. Browns included straight- or curly-haired 2's, also 3, 4, and 5. White included only rating 6. Two occupational levels were distinguished, high (ratings 1-3) and low (ratings 4-6).

Sex and type of school were considered as possible criteria. But an analysis of variance applied to the scores of 9-year-olds on the man test, with color, occupation and sex as the criteria, resulted in *F*'s significant for color and occupation but not for sex. For the 36 boys and 37 girls of this age the *F* was only .01. For several other groups and tests conventional *t*-test comparisons were made and no significant sex differences found.

Type of school was not used as a criterion for two reasons. First there were not enough black and brown pupils of high occupational level to provide all the sub-groups needed. Second, in cases where private-government com-

parisons could be made, *t*-tests showed a striking lack of significant differences. For example, comparing in the 9-year-group all 27 private school pupils with all 52 government school pupils on the man test, we find a government mean of 15.69 and a private mean of 13.41. But the difference, 2.28, is not significant, *t* being 1.64. Again on the blocks test, age 10, we have means of 6.123 and 6.67 respectively for the white private school pupils and the 9 white government school pupils. The difference of .547, with a *t* of .38, is quite non-significant. A third example we take within the white group, age 9 where there are 8 in private and 8 in government schools. Means on the blocks test are 5.25 and 5.85 respectively, with a non-significant *t* of .35.

The criteria used in the basic analysis of variance, then, were color, age, and occupation. For certain tests a fourth criterion, self-confidence rating, was employed.

2. Data

Table 3 summarizes the analysis of variance for the five junior tests and Table 4 gives the leading means and *SD*'s. Table 5 summarizes the analysis of the four senior performance tests and five senior "verbal" tests. Table 6 summarizes the analysis of variance with self-confidence as one criterion for senior opposites and for two junior tests, age 9 only.

TABLE 3
SUMMARY, ANALYSIS OF VARIANCE OF FIVE JUNIOR TESTS

Source	<i>df</i>	Man <i>MS</i>	<i>F</i>	<i>df</i>	Blocks <i>MS</i>	<i>F</i>	<i>df</i>	Purse <i>MS</i>	<i>F</i>
Color	2	11.44	.37	2	42.54	5.86**	2	8.727	1.36
Age	2	286.07	9.26**	2	51.85	7.14**	2	34.047	5.30**
Occ.	1	18.98	.61	1	72.05	9.92**	1	.033	.03/6.4
Error	164	30.88		163	7.257		163	6.413	
Pict. Class.				Squares					
Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>df</i>	<i>MS</i>	<i>F</i>			
Color	2	69.45	10.21**	2	6.674	2.15			
Age	2	9.06	1.49	2	34.655	11.18**			
Occ.	1	118.52	19.43**	1	7.329	2.36			
Error	161	6.07		163	3.101				

** Significant at 1 per cent.

D. DISCUSSION

The *draw-a-man test* elicited eager response. A few who required encouragement all did proceed and there were no zero scores. It is clear that the inclusion of the private schools and the children of the "best families" did not result in relatively higher scores for white groups, the *F*'s indicating no significant differences associated with either color or occupation. In Table

TABLE 4
JUNIOR TESTS, MEANS BY AGE, OCCUPATION, AND COLOR*

JUNIOR TESTS, MEANS BY AGE, OCCUPATION, AND COLOR																								
Age	Color	N			Man			Blocks			Purse			Pictorial Classification			Memory Squares							
		H	L	M	High Occ.	Low Occ.	SD	High Occ.	Low Occ.	SD	High Occ.	Low Occ.	SD	High Occ.	Low Occ.	SD	High Occ.	Low Occ.	SD					
8	White	11	8	12.8	6.7	13.75	2.2	5.18	3.0	1.86	1.9	4.20	2.4	3.38	2.3	6.20	2.4	4.43	2.7	3.41	2.1	1.57	1.3	
	Brown	8	10	13.25	5.7	12.4	1.2	4.3	2.7	3.8	1.4	4.88	2.3	4.80	2.4	6.38	1.7	4.70	2.7	2.06	1.0	2.15	1.2	
	Black	4	7	18.25	4.6	11.57	4.2	5.00	1.6	1.63	1.6	3.50	2.2	3.29	2.0	5.00	3.0	2.00	1.5	1.75	.8	1.29	1.5	
9	White	10	16	15.2	3.5	13.94	5.1	7.67	2.9	5.56	3.3	3.80	2.5	4.67	2.9	7.5	2.3	5.06	2.6	3.72	1.4	2.97	1.4	
	Brown	8	16	17.5	7.6	13.75	6.1	5.33	3.0	3.00	2.9	1.88	1.16	3.50	2.6	6.	2.7	4.69	2.4	4.38	2.3	3.16	1.3	
	Black	5	19	15.00	4.8	15.63	4.7	4.4	1.5	4.37	2.5	4.40	3.5	3.26	2.6	5.2	2.1	3.56	2.2	3.60	2.3	2.95	1.1	
10	White	6	8	19.67	10.7	18.13	5.6	8.17	3.9	6.13	2.7	5.33	3.1	6.87	2.0	7.17	2.3	6.13	2.3	4.33	1.6	4.25	.9	
	Brown	6	13	15.17	4.0	20.62	7.9	6.17	1.6	6.15	2.6	4.17	2.3	4.31	2.5	6.50	1.9	6.23	2.8	3.75	1.3	4.42	2.0	
	Black	4	12	19.00	7.2	16.75	5.3	3.50	1.5	4.42	2.9	6.75	.8	4.67	2.7	5.25	1.3	3.83	3.0	3.13	1.1	2.96	1.8	

* H = high occupational level, L = low occupational level.

TABLE 5
SUMMARY, ANALYSIS OF VARIANCE, NINE SENIOR TESTS

Source	df	Blocks		Forms		Mem. Sqs.		Pict. Class.			
		MS	F	MS	F	MS	F	MS	F		
Color	2	16.685	1.12	9.099	1.79	2.05	.43	1.89	.63		
Age	2	263.28	5.91**	30.254	5.94**	14.63	3.35*	6.69	2.45		
Occ.	1	9.57	.64	4.81	.93	4.36	.93	14.19	4.76**		
Error	164	14.84		5.094		4.67		2.98			
Source	df	Verbal Class.		Opposites		Numbers		Analogies		Reasoning	
		MS	F	MS	F	MS	F	MS	F	MS	F
Color	2	222.65	7.12**	649.41	6.61**	14.00	.75	273.46	4.62*	3.35	1.67
Age	2	322.97	10.33**	873.03	8.88**	118.54	6.33**	408.66	6.9**	29.45	6.9**
Occ.	1	278.68	8.91**	347.74	3.54*	2.05	.11	145.437	2.45	4.71	1.15
Error	164	31.25		98.24		18.73		59.146		4.26	

* Significant at 5%; ** significant at 1%.



TABLE 6
ANALYSIS OF VARIANCE OF SENIOR OPPOSITES, JR. BLOCKS, AND PICT. CLASS.

Source	df	Sr. Opposites		Age 9	Junior Blocks		Age 9	Jr. Pict. Class.	
		MS	F	df	MS	F	df	MS	F
Color	2	909.62	12.41**	2	31.94	7.6**	2	21.79	10.9**
Age	3	1007.31	13.74**	—	—	—	—	—	—
Occ.	1	349.46	4.76*	1	79.24	18.87**	1	71.98	36.01**
Conf.	1	3096.24	42.25**	1	74.67	17.78**	1	61.43	30.75**
Error	142	73.27		58	4.20		59	1.998	

* Significant at 5%; ** significant at 1%.

4, means for no one color group are regularly higher or lower than those for other color groups. Age is a very significant factor as is expected of a good test for this age-range.

The *blocks test* was very difficult for the juniors. There were 8 zero scores among the 48 8-year-olds, and 5 among the 73 9-year-olds, though there were none in the 10-year-group. Table 3 shows very significant *F*s for all three factors and Table 4 indicates a definite white superiority. In both occupational groups the whites are superior to both the browns and the blacks in 3 out of the 18 comparisons; and in all 18 comparisons white children are superior to the black. In certain of the black-white comparisons we found the differences were not significant. For example the 16 9-year-old white children of low occupation average 1.19 points more than the 19 black pupils but since *t* is only 1.64, the difference is not significant. Nor is the difference between white and black in the high occupation of age 9, for though the difference in means is 3.27, the *t* of 1.27 shows its lack of significance.

In year 10 we notice more striking color differences in scores in the blocks test, while in age 8 the brown pupils are superior to white ones of the same (low) occupational group. Taking all 9-year-old children regardless of occupation we find a white mean of 6.32 and a black mean of 4.38, and the difference, 1.94, is significant at the 5 per cent level of confidence.

Whatever is the nature of the differences in performance on the junior blocks, the whites clearly excel. Apparently faced with this difficult test, the black and brown children were less well able to deal with it than the white.

The seniors found the blocks test easier, and only three of 170 made zero scores. Among these seniors four were reported by the teacher to be mentally deficient. Three of these were black, in years 11, 13, and 15, and one was brown, in age 13. The *F* for age is highly significant for the senior blocks, but no differences associated with either color or occupation are

indicated. These clear-cut results suggest that the color differences for the juniors may have been related to the difficulty of the test for them, and that the junior difference may be due to some other factor than greater inability of the colored to deal with the spacial relationships involved.

For the juniors, distributions on the *geometrical forms test* were far from normal. Of the scores for age 8, 26 per cent were zero, 56 per cent either zero or one. Hence analysis of variance of junior scores is not suitable. For age 9 conventional color comparisons were made, though 48 per cent of the scores were zero or one. Means for white, brown, and black children are 1.92, 1.72, and 1.41 respectively. None of the differences are significant, t being 1.45 for the white-black difference and .61 for the white-brown difference.

Even for older pupils (seniors) the forms test proved hard and did not arouse as much interest as other tests. The analysis of variance (Table 5) showed age to be highly significant, but F 's for both occupation and color indicated that there were no significant differences associated with those variables.

On the *lost purse test*, given to juniors only, F 's for neither color nor occupation are significant but age is shown to be highly significant. The test evoked hearty response and some very sophisticated plans of search were presented.

The *memory squares test* elicited positively enthusiastic response from the junior pupils and almost as much from the seniors. Highly significant age differences among juniors are indicated, and a significance of age at the 5 per cent level among seniors. The increase in mean score with age is fairly consistent.

The F tests indicate that in the memory squares test no significant differences are associated with either occupation or color, in either junior or senior groups.

On the *pictorial classification test* the junior results show no regular increase in mean score with age, nor is the F for age at all significant. Highly significant differences are, however, associated both with occupation and with color. The black means (Table 4) are lower than the brown and the white ones in all of the six age-occupation groups, the brown means somewhat lower than the white in four of the six groups.

Conventional comparisons of color group means, age 9, using the t test, showed the difference of 2.09 between the means for 26 white pupils and 24 black pupils to be highly significant, t being 2.98. The brown mean was less by .91 than the white mean, but the t for this difference, 1.24, is not

significant. Similar though less marked differences are found in the 8 and 10 year groups.

Conflicting results in this test for juniors, and absence of increase with age, suggest that the test as given was too difficult for juniors. It proved definitely more difficult for black than for white pupils aged 8, 9, and 10.

Table 5 indicates that among the older pupils there were no significant color differences on the pictorial classification test—the black pupils did just as well as the white ones. Though this test did not show a significant effect of age, the lack of any significant color differences might lead one to suspect that the highly significant color F for the juniors may have been due to some other factor than inability on the part of the black pupils to understand and deal with the conceptual relationships involved.

If the hypothesis is sound that there are factors other than race that could account for the color differences found among some junior groups but not among senior groups on these performance tests, it is reasonable to suppose that inability to understand the spoken instructions, i.e., a language handicap, might be one of those factors. It is true that in spite of the simplicity of Cayman society, there was unquestionably a generally higher level of culture among the white than among the colored (brown) and black groups, regardless of occupation. If within either our high or low occupational groups we could rank people on a scale of class or social (cultural) differences, there is no doubt that the blacks would be at the bottom of the scale. The more cultivated Cayman families sent their children to private schools, and in the two best (and incidentally the two largest) of these schools any visitor would be struck by the greater social ease and the better vocabulary of the pupils as compared with that of most (but not all) government school pupils.

If the above analysis is sound we would expect that the black and brown pupils who equalled the whites in three of the five junior tests and in all of the four senior non-language tests, would do relatively much less well in verbal tests. We expected for example that the very dark seniors who equalled whites in the pictorial classification test might be inferior to them in verbal classification which presumably involves some similar modes of thinking. Results of the four verbal and numbers tests provide evidence bearing on this point.

The analysis of the *verbal classification test* yielded very significant F 's for color. The same very dark seniors who had done just as well as the white pupils in classifying geometrical figures, were definitely inferior to them in classifying words. When we consider that such rows of words were

used as "garage, laboratory, school, library" (No. 21), we are helped in understanding this difference.

Results of analysis of the *opposites test* were similar, except that they indicated a less close association of score with occupational rating. In this test the white mean (means not given here) was the largest for all age groups, and in only one, age 12, was the black mean as high as the brown. This verbal test included, along with word pairs such as "shy-timid," others such as "intensity-loudness," and, toward the end, "decadence-progress."

Results of analysis of variance of the *number completion test* indicate a high significance for age but no significant color or occupational differences. Comparing means for color groups on this test (not given here) we found the black means higher than either the white or the brown for ages 11 and 12, but lower than either for ages 13 and 14. No color differences are significant—the black children handle these number relationships as well as the white children.

The final test which we called "*reasoning*," consisted of two parts the scores on which are combined for this report. As in the numbers test the analysis indicated a high significance for age differences but none for occupation or for color. This result might seem to be in conflict with the findings for verbal classification, opposites and analogies, but it is true that the words used are not so academic as many in the latter tests. Also in this "reasoning test" even if a single word were not familiar, the meaning of the sentence could be got through the context. Moreover in the proverbs section we are dealing with a type of thinking common among all classes in these islands.

In general then we have a situation where the only clear color differences that occur are in tests which (a) seem to be too difficult for the ages tested or (b) use language which would be likely to be fully understood only in homes of relatively superior culture. Is it possible to isolate any factor which, both in the two difficult junior tests and in the senior language tests, might reasonably account for the inferiority on these tests of the darker pupils?

E. ANALYSIS WITH SELF-CONFIDENCE RATING AS A CRITERION

For certain tests we made an analysis of variance using as one of the criteria teachers' ratings on self-confidence. This rating, we had noticed, was frequently different from a teacher's other ratings of the same pupil. One difficulty in making this analysis was that there were too few black pupils of low occupational level who rated high in self-confidence, to provide cases for all sub-groups. This was dealt with by going back to the original

record sheets to find pupils rated 4 in occupational level, where it had been noted that the case was borderline and that a 3 rating (high) might have been given. Changing the occupational rating accordingly, we had at least one in each of the groups.

Table 6 summarizes this special analysis of variance for seniors on opposites, and for juniors age 9 on blocks and pictorial classification. This table does suggest the importance of self-confidence as a factor in performance on the opposites test. F 's for age and color are still highly significant but the F for confidence is more than three times that for color. Low confidence means (not given here) are in most cases less than high confidence means, in all color-age groups. A similar situation is reflected in the F 's for junior blocks and pictorial classification. Whatever it is that self-confidence ratings may be based on affects performance on the tests. We have too few cases in certain groups to justify further analysis of these particular results, but we suggest that some measure of self-confidence should be used in studying racial differences in intelligence. On the island of Grand Cayman as in Jamaica, respect for one's betters has long been taught to children as one of the virtues. And on this island, more than in Jamaica, a white skin is highly respected. It would be natural then that when pupils in a school perform a task that is very hard for all of them, the black skinned ones would be more likely to work with doubt and with misgivings than white pupils of the same native ability, and thus do less well.

F. SUMMARY

In this study the commonly found black inferiority on intelligence tests occurred only in tests for younger children which did not show the advance in age expected of a good test, and in the most academic tests for older children. On three out of five performance tests there were no significant color differences among juniors. The same dark-colored seniors who were inferior to whites in verbal classification, opposites, and analogies, did fully as well in all the non-language tests, including pictorial classification.

The older colored pupils also did as well as the whites in the test involving number relationships and in the final "reasoning test" which called for certain symbolic operations such as characterize the thinking we call reasoning. It is significant that this last test did not employ very uncommon words and did deal with subject matter, such as family relationships and proverbs, with which colored pupils were probably as familiar as the more cultivated white pupils. The present study, then, does not lend support to the conclusion that colored inferiority in intelligence tests has a racial basis.

REFERENCE

1. SHUEY, A. M. The Testing of Negro Intelligence. (Foreword by H. E. Garrett.) Lynchburg, Va.: Randolph-Macon Woman's College, 1958. Pp. 351.

3516 Tallyho Lane

Madison 5, Wisconsin

A NEW METHOD FOR THE INVESTIGATION OF THERMAL SENSITIVITY*¹

Department of Psychology, Florida State University

D. R. KENSHALO, JOHN PAUL NAFE, AND W. W. DAWSON

A. PROBLEM

Most of the investigations of thermal sensitivity do not agree upon the magnitude of the thresholds reported. They have been found to be as high as 11.5° C. (Lele, 1954) and as low as 0.003° C. (Hardy and Oppel, 1936). Furthermore, the individual responses have been so variable that at least one author concluded that: "Attempts to measure the sensory threshold for warmth with radiant stimuli reveal an almost bewildering variability of responses" (Wright, 1955, p. 342).

We believe that much of the inter- and intrasubject variability of the thermal threshold measures encountered by investigators results from continual changes in skin temperature (Sheard, 1944). For example, if the experimenter measures skin temperature to determine the range of the test stimuli, by the time a threshold has been obtained skin temperature may have changed by 0.2° C. or more, thereby vitiating the original measure. Skin temperature may have shifted by an even greater amount on subsequent threshold determinations. At least three investigators (Culler, 1926; Ebaugh and Thauer, 1950; Lele, 1954) have noted large variations in thermal thresholds with systematic changes in skin temperature.

Apart from the above consideration, comparison of results obtained in various experiments is made difficult because of the varying conditions of measurement. At least two of these conditions are the size of the area stimulated (Hardy and Oppel, 1937; Hergert, Granath, and Hardy, 1941; Hergert and Hardy, 1942; Lele, Weddell, and Williams, 1954), and the rate of change of the stimulating temperature (Hensel, 1952; Lele *et al.*, 1954). Other conditions of possible importance, which have not been dealt with, are the part of the body stimulated; e.g., forehead vs. forearm vs. hand, and the age of the subject.

In this paper a new apparatus is described which controls many of the

* Received in the Editorial Office on September 10, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The work reported here is supported by Contract No. DA 49-007-MD-683 between the Office of the Surgeon General, Department of the Army, and Florida State University. Reproduction in part or in whole is permitted for purposes of the United States Government.

variables known to affect the thermal thresholds and which allows a systematic investigation of several of these variables. Preliminary data are presented on the threshold of warmth and cool, as a function of the temperature to which the skin is adapted, to demonstrate the accomplishments and limitations of the equipment.

B. METHOD

1. *Apparatus*

The two principal means of delivering thermal energy to the skin are by radiation and by conduction. The former has been widely used, especially by Hardy and Oppel (1936, 1937), and has several advantages, all of which have to do with the control of the strength of the stimulus, its area and duration of application, and its measurement. It suffers from at least two serious limitations. First, there is no means of presenting cooling stimuli with the ease and control of warming stimuli, and second, it is questionable that a constant skin temperature can be maintained between stimulus presentations. A third drawback is that an appreciable time is required to attain a given change in skin temperature by means of radiation. One can either warm the skin very gradually until *S* reports a thermal sensation, in which case the data are confounded by the adaptation process (Hahn, 1949), or *E* may present a strong radiation for a short period and ask *S* to indicate when he first feels a thermal sensation. In the latter instance the amount of radiation delivered is confounded by the reaction time of *S*, a measure known to be variable, especially at threshold (Lele and Sinclair, 1955). For these reasons we chose to design a device which delivers its thermal stimuli by conduction, the stimuli being obtained by circulating water of controlled and measured temperatures.

a. Stimulus Delivery Apparatus. The thermulator,² shown in Figure 1, is the part of the apparatus which delivers a temperature change (ΔT) to the skin. It is similar in design to that described by Hensel, Ström, and Zotterman (1951). It consists of a chamber (1) formed from a block of lucite 3.6 x 2.6 x 2.0 cm. deep. The bottom of the chamber is a silver plate 2.0 x 3.0 x 0.0254 cm. thick, cemented to the lucite, and is the element which contacts the skin. Two water jets (2) are directed into the chamber. A scoop (3) mounted just above the chamber is of such width that upon rotation it interrupts either jet before the stream of water reaches the lower chamber. By rapidly switching from one stream to the other a stimulus of the magnitude of the difference between the two jets (ΔT) can be presented. Measurement of the ΔT at the bottom of the silver plate showed 63.2 per cent of the ΔT took place in approximately 0.4 sec. (see Figure 2).

² Our thanks to Dr. K. S. Wagoner who made the thermulator.

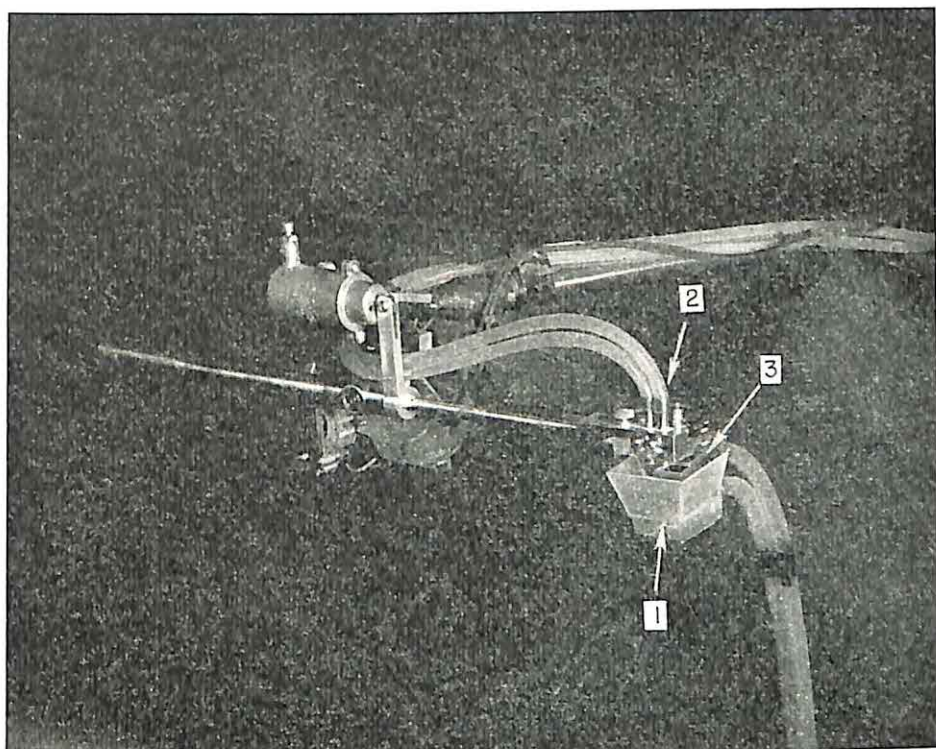


FIGURE 1
THE THERMULATOR

The chamber, which rests on the skin, is shown as (1). Either of the two water jets (2) may be interrupted by rotation of the scoop (3).

The scoop is rotated by a solenoid activated by a modified Hunter timer (Hunter, 1949). The timer is set so that upon activation of a switch by *E*, the scoop rotates and presents the ΔT for 10 sec., or until *S* closes a key indicating that he has sensed a temperature change, whichever is sooner, after which the scoop returns to its original position.

During the course of the threshold measurements the temperature of the waters flowing into the scoop and the chamber are continuously monitored by thermistors mounted in the center of the bottom of the scoop and chamber. These are each sections of Wheatstone bridges and are calibrated so that the scoop and chamber bottom temperatures are registered on a galvanometer. These can be read directly to 0.1°C ., and estimation to the nearest 0.025°C . is possible.

The thermulator is mounted on a counterbalanced arm so that the weight

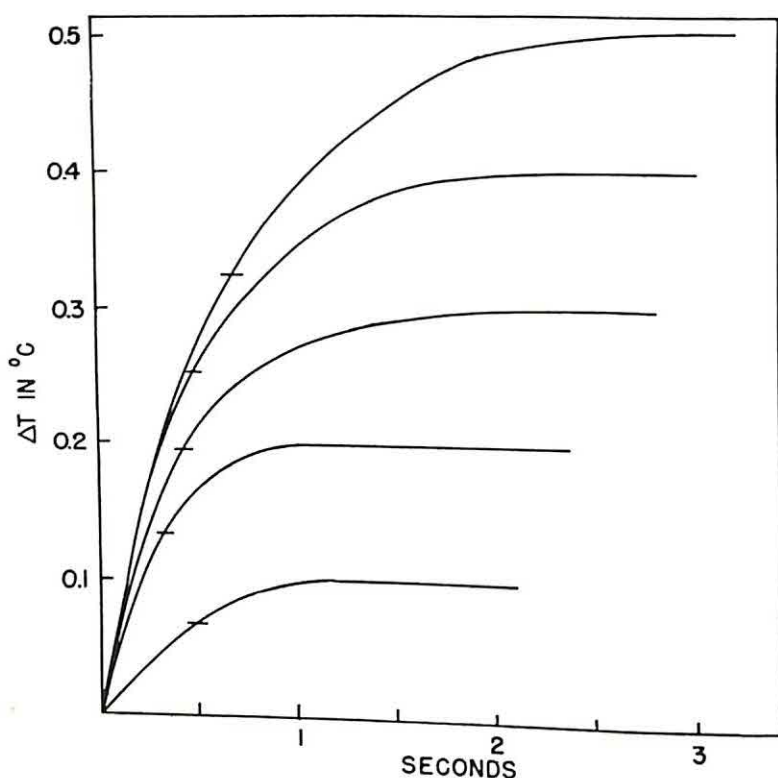


FIGURE 2

THERMULATOR RESPONSE TO STEP-FUNCTION TEMPERATURE CHANGES

The temperature changes on the bottom of the chamber plate upon switching from one jet to the other. The horizontal marks on each curve are at the point where 63.2 per cent of the total change in temperature has occurred.

with which the thermulator rests on the skin can be adjusted up to a pressure of 25 grams per sq. cm.

b. Stimulus Control Apparatus. Water for the thermulator jets is provided from six 4-gallon and one 8-gallon thermostatically controlled stainless steel tanks, as shown in Figure 3. The sensing element of the thermostat (1) is a 34A3 Victory Engineering thermistor encased in a stainless steel tube. The thermistor constitutes one section of a *DC* Wheatstone bridge. Imbalance of the bridge is sensed by a differential modulated amplifier. After demodulation the output of the amplifier operates a thyatron controlled relay. The thermostat controls the temperature of the water in the tanks to less than $\pm 0.001^{\circ}\text{C}$., as measured by a Beckmann differential thermometer. With this high degree of sensitivity the thermostat would "chatter" due to convection currents in the tank were it not for a large filter condenser in the demodulator section on the output of the amplifier.

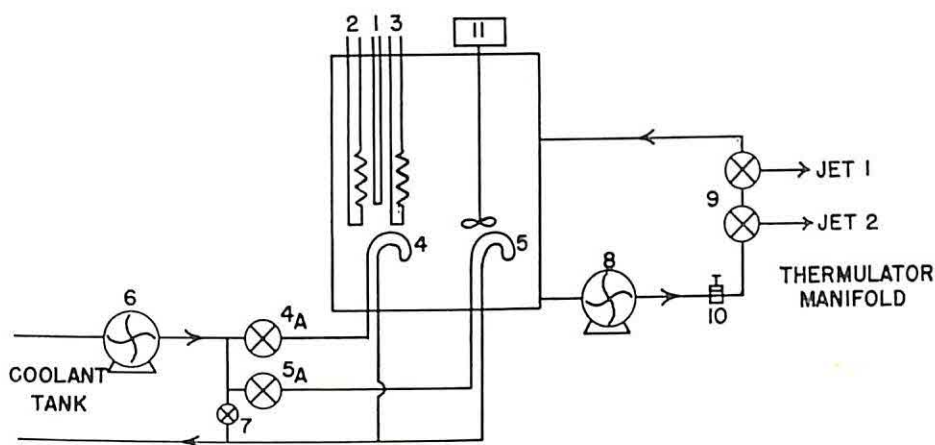


FIGURE 3

SCHEMATIC DIAGRAM OF HEATING, COOLING, AND CIRCULATING SYSTEMS IN EACH OF THE SEVEN TANKS

The thermostat sensing element is shown at (1). The heaters (2 & 3) are 75-watt display case bulbs and a Chromlox 330-watt immersion heater. High and low capacity coolers (4 & 5) are two feet of one-fourth inch and four feet of one-half inch copper tubing. The flow of coolant through these is controlled by magnetic valves (4A & 5A). A positive displacement pump (6) circulates coolant through the pressure relief valve (7) and back to the refrigerated tank. Water in the temperature controlled tank is kept in motion by an agitator (11). Water from the tank is pumped by a centrifugal pump (8) through the two Circle Seal 2-way valves (9) and back to the tank. Switching either valve directs water into either of two manifolds connecting similar valves on the other tanks to the two water jets on the thermulator.

Heating and cooling of the tanks is accomplished by two heaters and two coolers. Each tank has a high capacity heater (3) and cooler (4). These are suitable for producing fairly rapid changes in the water temperature, but have too large a thermal inertia for the maximum sensitivity of the thermostats to be realized; hence, in order to maintain the temperature set by the thermostats, each tank is equipped with a low capacity heater (2) and cooler (5). Coolant (a mixture of glycol and water) at -5°C . is circulated by the pump (6) from a 50-gallon refrigerated tank through a pressure relief valve (7) and back to the refrigerated tank. Magnetic valves (4A & 5A), when opened, allow coolant to circulate through cooling coils in the tanks. Both the heaters and magnetic valves of the coolers are operated through appropriate relay switching circuits by the relays on each thermostat. The circuit is such that E can select heater only, cooler only, or oppose the heaters and coolers. The last condition is particularly useful when the tank temperatures are set at about that of the room. The switching is designed so that when the apparatus is turned on the high capacity elements are operational but once the desired tank temperatures are attained, the high

capacity elements are automatically deenergized and the low capacity elements energized.

Each tank is equipped with a pump (8) which circulates the water of proper temperature through manually operated valves (9) and back to the tank. When the valve (9) is opened water is diverted to a manifold connecting it to the output of similar valves supplied by the other six tanks, and thence to the jets on the thermulator. In this way *E* can divert water from any tank into either of the jets on the thermulator. The rate of flow is adjusted by the screw clamp (10). The entire system (valves, pumps, and tubing) up to the thermulator reaches thermal equilibrium if the apparatus is activated approximately an hour before the experiment session commences.

2. Procedure

a. Subject Controls. All of the equipment is located in an air-conditioned, 12 ft. by 12 ft. room. The temperature of the room is maintained at $22 \pm 1^\circ \text{C}$. The humidity varies between 40 per cent and 60 per cent. Figure 4 shows the entire apparatus ready for use. The *S* is seated in the adjustable chair in the foreground, the right arm of which has been covered with foam rubber. Signal switches activated by *S* are located at the end of the left arm. The *E* is seated on a stool in back of, and to the right of *S*, and in front of the relay rack on the right. This rack contains (from top to bottom) the switching circuits controlling the thermulator scoop (2) meters and range switches for monitoring the temperature in the thermulator and scoop (3), and the panel containing the 2-way valves (4 and 9 in Figure 3). The left relay rack contains the seven thermostats, while to the left of this is the refrigerated tank. Just behind the relay racks the stand containing the water tanks, pumps (8 in Figure 3) and magnetic valves (4A & 5A in Figure 3) can be seen.

b. Procedural Controls. Since the procedure will differ somewhat, depending upon the variable under investigation, the procedure employed in the investigation of the rôle of skin temperature upon the warmth and cool threshold is presented as a model in discussing these controls.

The *S* was seated in the adjustable chair and allowed to become acclimated to the room temperature for 30 minutes. Opaque goggles covered his eyes to prevent him from observing the manipulations of the apparatus by *E*. The area of the body where the thermal thresholds were measured in this preliminary report was the dorsal surface of the forearm approximately three inches below the bend of the elbow. This site had been shaved 24 hours before measurements were started. The thermulator was placed on the

site and rested there with a weight of 16 grams per sq. cm. The large water tank (8-gallon) had previously been set at one of the six adapting temperatures (28°C. , 31°C. , 33.4°C. , 36°C. , 39°C. , and 42°C.), and was now allowed to flow from one of the two jets into the chamber of the

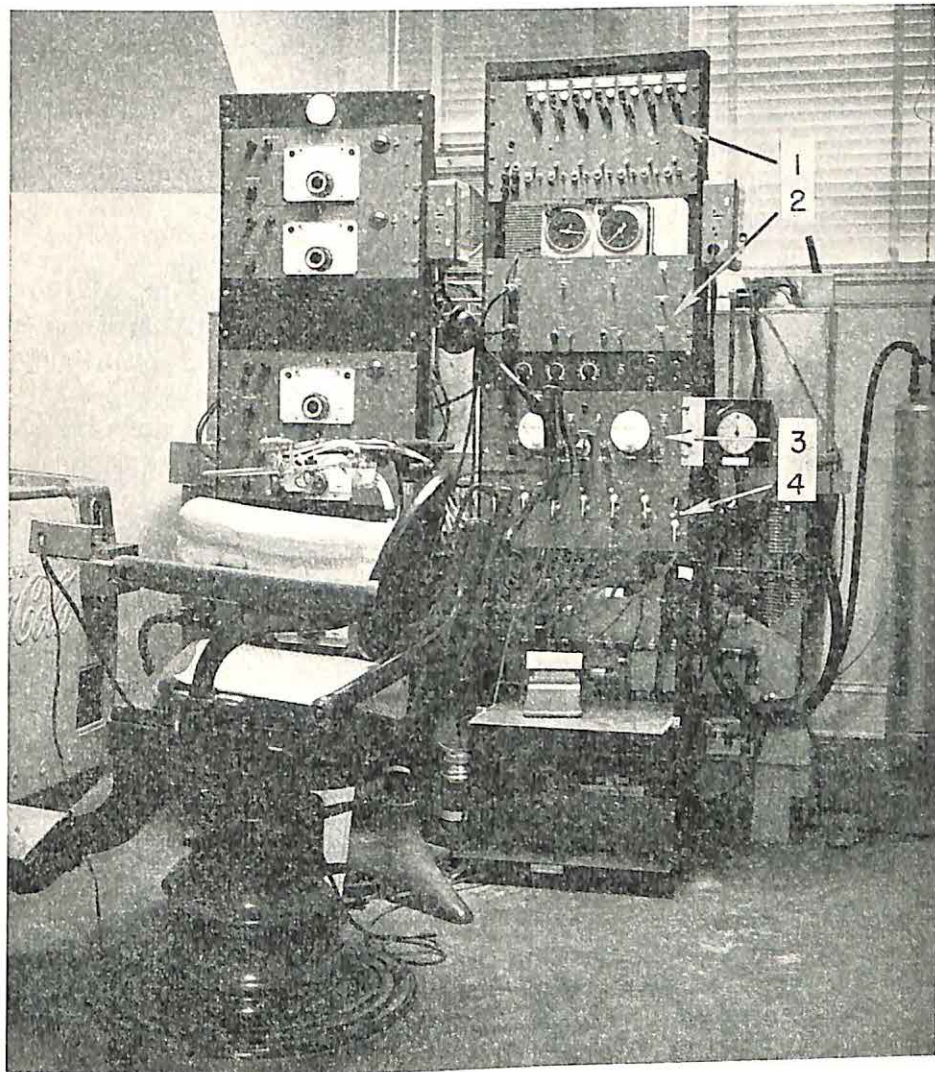


FIGURE 4
THE COMPLETE APPARATUS

The right relay rack contains the switching circuits (1) of the thermostats (left relay rack); clocks and control panel for the thermulator are indicated as (2); meters for monitoring the temperature of water in the scoop and chamber are at (3), while (4) shows the 2-way valves which determine which tank supplies the scoop and chamber.

thermulator. The *S* was allowed to adapt completely (disappearance of the thermal sensation to this temperature). The time allowed for adaptation ranged from a minimum of 20 minutes for the temperatures close to that of the skin, to 45 minutes for the two extreme temperatures. Complete adaptation within 45 minutes could not be achieved either above or below this range, therefore, these data are limited to these adapting temperatures. The other six tanks had been set so that water was delivered at the thermulator in increments or decrements of 0.1°C ., depending upon whether the warm or cool threshold was being measured. Exception was made in the case of the warm threshold at an adapting temperature of 42°C . where increments of 0.05°C . were used.

Both the ascending and descending series of the psychophysical Method of Limits (Guilford, 1954) were used in determining each threshold. Measurement of the warm threshold at 36°C . may serve as an example for the exact procedure followed at each of the adapting temperatures in obtaining the warmth and cool thresholds. Preliminary trials at the adapting temperature of 36°C . indicated that the warm threshold would be at a ΔT of about 0.3°C ., therefore, the tanks were set to deliver water at the thermulator at 36°C . (the adapting temperature), 36.1°C ., 36.2°C ., 36.3°C ., 36.4°C ., 36.5°C ., and 36.6°C . When *S* reported that adaptation was complete, additional time was given to complete the 20 min. minimum adaptation time. The *E* started water of 36.1°C . into the scoop. A "ready" signal was given, followed in 5 sec. by the ΔT . The ΔT remained for 10 sec., after which the adapting temperature was again presented. The *S* had been instructed to push a button on the left arm of the chair if he was aware of a temperature change in the chamber. Pushing the button caused the thermulator to return to the adapting temperature before the 10 sec. period had elapsed, and also indicated to *E* that *S* had felt the temperature change. The next higher ΔT was presented 1 min. after the first had been presented. This allowed a minimum of 50 sec. for readaptation to 36°C . between each presentation of a ΔT . The descending series followed the same procedure except that successively lower temperatures were presented until the subject no longer reported a change in temperature. A total of 20 series, 10 ascending and 10 descending, were run for both the warm and cold threshold at each of the adapting temperatures. Since six ΔT s were presented in each series each threshold is the result of 120 judgments. The temperature midway between those at which *S* gave consistent reports was taken as the threshold of that series.

One of us (Kenshalo) acted as a subject, while another (Dawson) was the experimenter. The *S* was a white male 35 years old at the time of the

experiment and thoroughly practiced as a subject in psychophysical measurements. While *S* was familiar with the nature of the experiment he did not know what temperatures were to be used at any given session, nor did he know the order of presentation of the ΔT s.

C. RESULTS

The data presented in Figure 5 should be regarded as preliminary, since only one *S* was used. They are presented here in order to facilitate discussion of the merits and limitations of the apparatus.

Each point on both the warm and cool threshold curves shown in Figure 5 is a mean of 20 measurements of the threshold. Reductions in temperature of approximately 0.15°C . at the adapting temperatures of 28°C ., 31°C ., and 33.4°C . gave sensations of cool. At adapting temperatures above these, the threshold increased to a point where decrements of approximately 0.65°C .

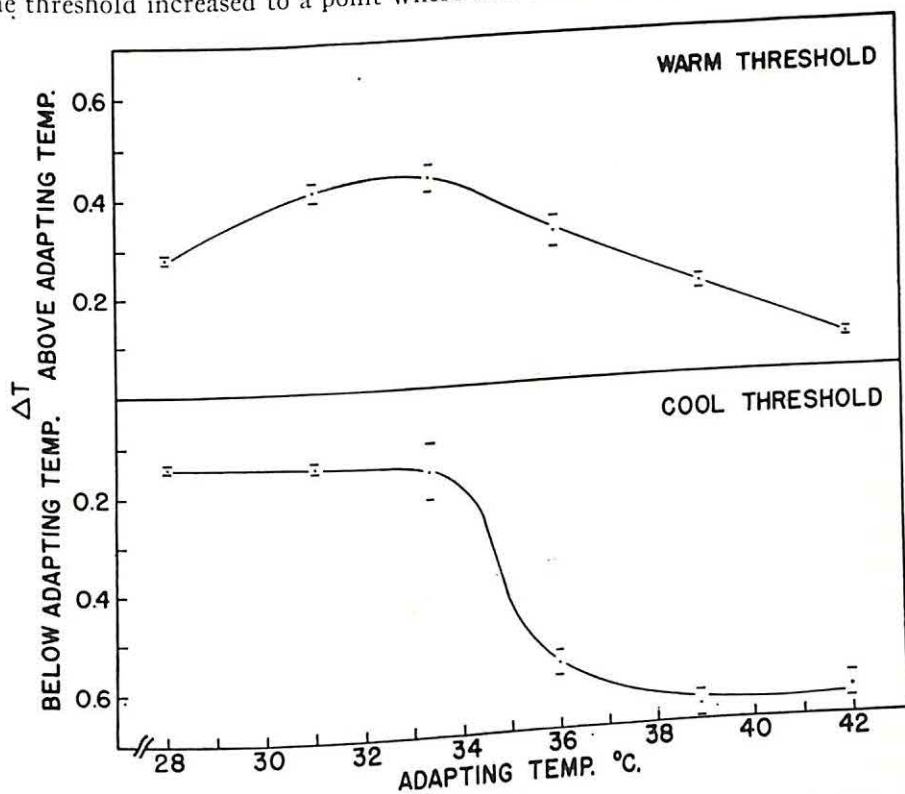


FIGURE 5
THRESHOLDS OF THERMAL SENSATION AT VARIOUS ADAPTING TEMPERATURES
Along the ordinate is shown the extent of the change in temperature from each adapting temperature necessary to produce a thermal sensation. The range marks at each point indicate the size of the standard error of the mean.

were required in order to be felt as cool. The warm threshold shows much more gentle changes. It rises from approximately 0.28°C. to 0.43°C. for adapting temperatures between 28°C. and 33.4°C. At higher adapting temperatures it gradually decreased until at 42°C. an increment of only 0.08°C. produced a report of warmth. The variability of the threshold measures at each adapting temperature for both the warm and cool thresholds are shown by the standard error of the means indicated by the horizontal lines at each point in Figure 5.

The character of the sensations produced by warm and cool threshold stimuli were quite different. The appearance of the cool sensation was quite abrupt. If it appeared at all there was little doubt on the part of *S* as to the character of the sensation. The warm, on the other hand, seemed to "build up" gradually. At ΔT 's very close to threshold it was difficult for him to determine if he felt a slight warmth or not. Similar observations were made by Goldscheider (1917).

D. DISCUSSION

The threshold curve for cool sensations, shown in Figure 5, is quite similar to that reported by Ebaugh and Thauer (1950) obtained by radiation methods. Their data, however, appear to be too variable to show any relationship between the warm threshold and skin temperature. The warm threshold curve of Figure 5 is the first report of such a relationship.

While the two curves shown in Figure 5 are not alike, in some respects they are similar. The warm threshold is large for low adapting temperatures, while that for cool is small. As the adapting temperature increases their position is reversed until at 42°C. the warm threshold is small while the cool threshold is large. Furthermore, their inflection points appear to occur at about the same adapting temperature. The shift in the warm threshold is not as abrupt as that occurring between 33.4°C. and 36°C. in the cool threshold curve.

There seems to be no reason to assume that the sensitivity of the thermal receptors is altered by changing the temperature of the tissue around them. We favor an explanation based on the work of Hertzman (1953) which indicates that the thermal conductance of the skin is altered by changing its temperature.

The temperatures reported in Figure 5 are threshold only insofar as they are the minimum ΔT , as measured in the chamber, capable of eliciting a report from *S*. They can not be construed as the minimum ΔT required to excite the thermal receptor. The same line of reasoning applies to the adapting temperatures reported. Attempts to measure the skin temperature

by means of a 40 B & S gauge copper-constantan thermocouple approximately 0.05 mm. beneath the cornium showed that a ΔT of almost 0.3°C . in the chamber was necessary to produce a change of only 0.1°C . at the thermocouple after 10 sec. Likewise, similar measurements of the effect of an adapting temperature of 6°C . above the initial skin temperature raised the skin temperature only 4.75°C . after 3 min. The high thermal inertia of skin (Hendler, Crosbie, and Hardy, 1958) as compared to that of the water and silver plate probably accounts for the major portion of this effect.

While each point on the curves assumes meaning only in relation to the other points on the curve and those obtained on subsequent experiments with this or similar apparatus, the situation is not unlike that in vision. Hecht, Schlaer, and Pirenne (1942) have reported that 5 to 15 quanta are necessary to excite a rod, yet 50 to 150 quanta must be delivered to the cornea.

The fact that the thermulator rests on the skin continuously from the beginning of the experiment to the end, circumvents the criticism that thermal threshold judgments are difficult to make when confounded with tactile stimuli as occurs when heated or cooled cylinders are applied to skin for each judgment (Woodworth and Schlosburg, 1954, p. 285). By allowing the thermulator to rest on the skin continuously tactile sensations have completely adapted by the time the thermal threshold measures were started. The *S* experienced only an infrequent, just detectable "bump" when the scoop was switched from one water jet to another.

Aside from this advantage, the fact that the thermulator can be maintained at a given temperature for long periods of time causes it to reduce variations in skin temperature as well as allowing the investigator to produce large changes in its temperature at will.

For the most exact measure of sensory functions, it is highly desirable that the change in energy level of the stimulus be accomplished as rapidly as possible. While the 0.4 sec. time constant required for a ΔT in this apparatus is much too slow to be ideal, it is considerably faster than warming the skin by radiant methods. Furthermore, the extent of the thermulator temperature change can be limited independently of the duration of application. With radiant methods the temperature change of the skin is a function of the stimulus strength and the duration of application. Variations in the strength-duration factor have yielded varying reports in the size of the thermal threshold (Lele, Weddell, and Williams, 1954; Ebaugh and Thauer, 1950; Hardy and Oppel, 1937).

E. SUMMARY

A new apparatus for use in the investigation of thermal sensitivity is described and preliminary data presented on the relationship between the thresholds of warmth and cool and the temperature to which the skin is adapted. It utilizes thermal conduction to deliver temperature changes to the skin, thereby allowing equal control in presenting both warm and cool stimuli. The stimulator remains in contact with the area under investigation at all times so that presentations of thermal stimuli are not confounded by tactile stimuli, and improved control of skin temperature is possible at all times. Temperature changes of as little as 0.05°C . can be reliably presented and provision is made to monitor these continuously. The preliminary data presented show that at adapting temperatures of 28°C . the warm threshold is relatively high, while the cool threshold is low. Between adapting temperatures of 33.4°C . and 35°C . both thresholds shift, the warm more gradually, so that at an adapting temperature of 42°C . the warm threshold is low while that for cool stimuli is large.

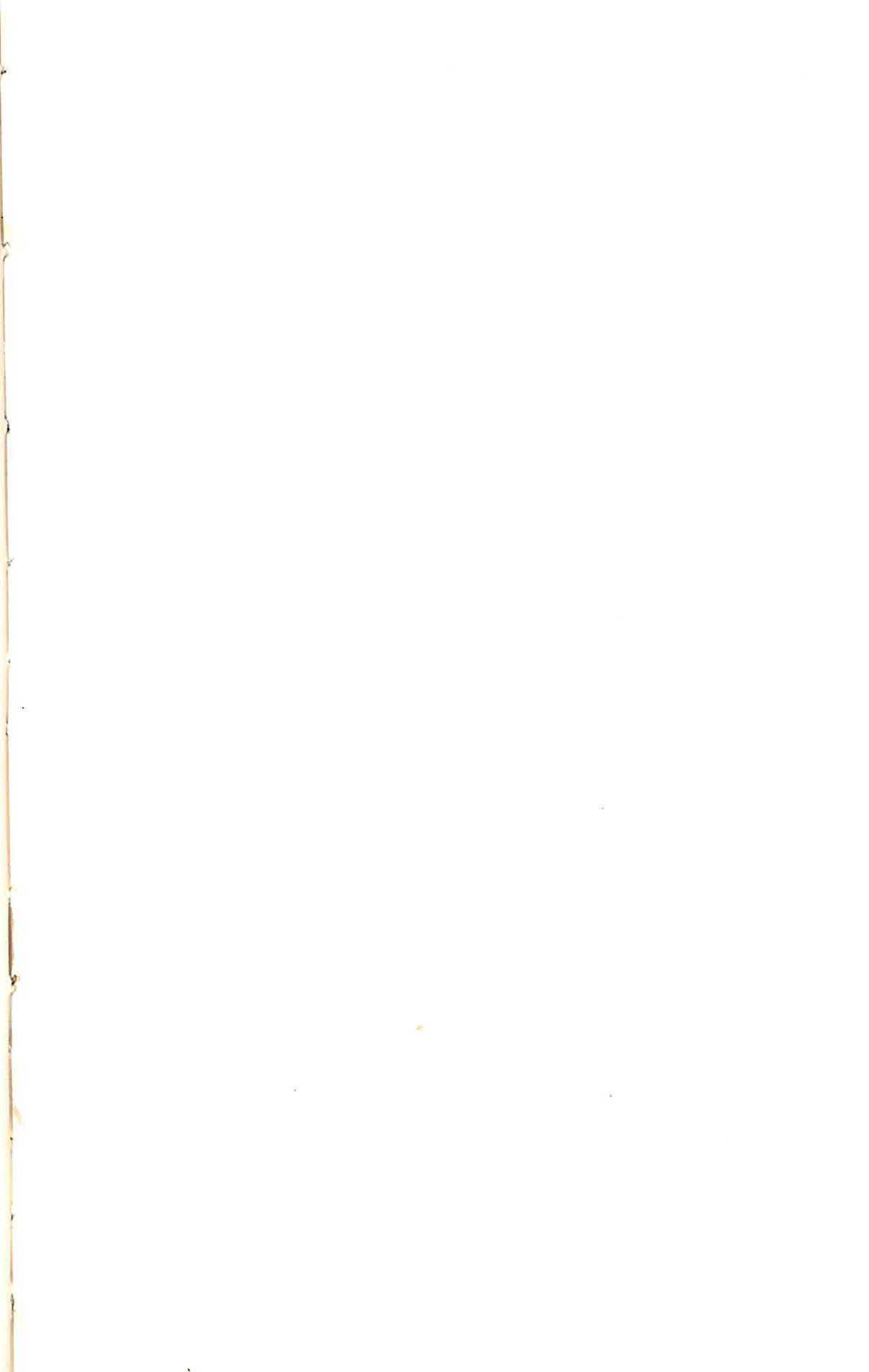
It is suggested that rather than a changing receptor sensitivity these data are explained by changes in the thermal conductivity of the skin.

REFERENCES

1. CULLER, E. A. Thermal discrimination and Weber's law. *Arch. of Psychol.*, 1926, **13**, 1-134.
2. EBAUGH, F., & THAUER, R. Influence of various environmental temperatures on the cold and warmth thresholds. *J. Appl. Physiol.*, 1950, **3**, 173-182.
3. GOLDSCHIEDER, A. Weitere mitterlungen zun physiologic der Sinnesnerven der Haut. *Pflüg. Arch. f. d. ges. Physiol.*, 1917, **168**, 36-88.
4. GUILFORD, J. P. *Psychometric Methods*, (2nd. ed.). New York: McGraw-Hill, 1954.
5. HAHN, H. *Beitrage zur Reizphysiologie*. Heidelberg: Scherer Verlag, 1949.
6. HARDY, J. D., & OPPEL, T. W. The thermal response of skin to radiation. *Physics*, 1936, **7**, 466-479.
7. HARDY, J. D., & OPPEL, T. W. Studies in temperature sensation: III. The sensitivity of the body to heat and spatial summation of the end organ responses. *J. Clin. Invest.*, 1937, **16**, 533-540.
8. HECHT, S., SCHLAER, S., & PIRENNE, M. H. Energy, quanta, and vision. *J. Gen. Physiol.*, 1942, **25**, 819-840.
9. HENDLER, E., CROSBIE, R., & HARDY, J. D. Measurement of heating of skin during exposure to infrared radiation. *J. Appl. Physiol.*, 1958, **12**, 177-185.
10. HENSEL, H. Physiologie der thermorection. *Ergebn. d. Physiol.*, 1952, **47**, 165-368.
11. HENSEL, H., STRÖM, L., & ZOTTERMAN, Y. Electrophysiological measurements of depth of thermoreceptors. *J. Neurophysiol.*, 1951, **14**, 423-429.
12. HERGERT, C. M., GRANATH, L. P., & HARDY, J. D. Warmth sense in relation to skin area stimulated. *Amer. J. Physiol.*, 1941, **135**, 20-26.
13. HERGERT, C., & HARDY, J. D. Spatial summation of heat. *Amer. J. Physiol.*, 1942, **135**, 426-429.

14. HERTZMAN, A. B. Some relations between skin temperature and blood flow. *Amer. J. Phys. Med.*, 1953, **32**, 233-251.
15. LELE, P. P. Relationship between cutaneous thermal thresholds, skin temperature and cross-sectional area of the stimulus. *J. Physiol.*, 1954, **126**, 191-205.
16. LELE, P. P., & SINCLAIR, D. C. Observations on the reaction time to cutaneous thermal stimuli. *J. Neurol. Neurosurg. & Psychiat.*, 1955, **18**, 120-124.
17. LELE, P. P., WEDDELL, G., & WILLIAMS, C. M. The relationship between heat transfer, skin temperature and cutaneous sensibility. *J. Physiol.*, 1954, **126**, 206-234.
18. SHEARD, C. Temperature of skin and thermal regulation of the body. In Otto Glasser (Ed.), *Medical Physics*. Chicago: Year Book Publishers, 1944. (Pp. 1523-1555.)
19. WRIGHT, G. H. The variability of the threshold of sensations of warmth evoked by radiation. *J. Physiol.*, 1955, **128**, 333-343.
20. WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental Psychology* (rev. ed.). New York: Holt, 1954.

Department of Psychology
Florida State University
Tallahassee, Florida





SCOPE OF INDUSTRIAL PSYCHOLOGY REFERENCES IN THE *ANNUAL REVIEW*, 1950-1959*

Human Relations Research Foundation, St. Louis, Missouri

H. MELTZER

A. INTRODUCTION

The intellectual craftsman, as described by Mills in *The Sociological Imagination*, learns "... 'how to get down to facts' and yet not get overwhelmed by them; how to anchor ideas to facts but not to sink the ideas" (14, p. 125). By so doing he arrives at a way of working with ideas which enlarge "... the scope of his sensibility, the precision of his references, the depth of his reasoning" (p. 120). Intended as a prescription for research workers in the field of social sciences, it is also, and even more so, an excellent prescription for the contributors to the *Annual Review*.

At the outset the editors of the *Annual Review* invited the contributors to be evaluative and interpretive, as well as merely to cover materials. How the various contributors to the approximately 18 chapters covered in the *Annual Review* have responded to this invitation would be interesting to know now that the *Annual Review* has been published for 10 years. A glance at the 10 volumes of the *Annual Review* will, of course, reveal what is to be expected, namely, that there are large personality differences in the manner in which the various contributors to the different areas have responded to this invitation.

Considering the fact that the *Annual Review* is being used with increasing frequency as an aid for orientation in the field by an increasing number of students and research workers in each field, an evaluation of the adequacy of the contributions so used takes on added significance. If these volumes are to serve to give an adequate picture of what has gone before, as well as of work currently in progress, it is important that the contributors make every effort to attain this objective.

B. THE PRESENT STUDY

In the present study an attempt is made to obtain a picture of the scope of reading materials used by the contributors to the nine chapters on industrial psychology in the *Annual Review* for the purpose of discovering the range, breadth, and depth of materials used in preparation for the writing of

* Received in the Editorial Office on September 11, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

the chapters. An evaluative study of the adequacy of the materials covered, considered in terms of desirable characteristics, is reported in a separate article (13A). Here the concern is with the extent to which the contributions show evidence of being sufficiently soaked in the literature of materials covered. The goal is for optimum selectiveness and coverage. It is possible to be oversoaked as well as unread. Mills, for example, warns of the danger of being oversoaked in literature because, says he, one may drown in it like Mortimer Adler did. The point is to know when you ought to read and when you ought not to read. How to be soaked in the literature so that none of the relevant and really significant is omitted and how not to get lost in details so that justifiable generalizing is inhibited is the guide for determining the amount and kind of source material to use. The latter danger is well expressed by Max Horkheimer: "The constant warning against premature conclusions and foggy generalities implies, unless properly qualified, a possible taboo against all thinking. If every thought has to be held in abeyance until it has been completely corroborated, no basic approach seems possible and we would limit ourselves to the level of mere symptoms" (14, pp. 122-123).

Put in more specific terms, the problems set for exploration include such questions as: Have the contributors been lost in gathering the facts? How broad has their outlook been? What kind of reading have they done in preparing for writing the chapter? Have they limited themselves to articles covered in the *Psychological Abstracts* for each year? Have they limited themselves to articles covered in easily available psychological journals? Have they made use of chapters in the *Annual Review*, other than their own specialty, which might be relevant? *Industrial psychology happens to be a field that does have to pick knowledge from other sources. In addition to references from strictly psychological journals, do they show enough breadth to include references indicating interest in and knowledge of related fields, for example, economics, cultural anthropology, and sociology?*

In short, to get some idea of the extent to which the contributors to the *Annual Review* have been too lost in details or have over-generalized, all references in the industrial psychology chapter in each volume have been tabulated. All references were tabulated both in terms of frequency as well as distribution of use by each of the contributors. From a study of the references used and manner of use, it is possible to obtain a picture of the nature of sources and resources used in working up the chapter as far as reading material is concerned.

C. INDUSTRIAL PSYCHOLOGY REFERENCES IN THE *Annual Review*, 1950-1959

Though the Editorial Committee of the *Annual Review* is still much concerned with the problem of topic structure and arrangement and expects to reconsider the philosophy which, so far, has governed the manner in which the *Annual Review* has split the area of psychology into approximately 18 sub-areas, industrial psychology as such has not suffered and apparently will not suffer from the changes that emerged from these reconsiderations. Industrial psychology has appeared as a chapter in every volume from the beginning until 1959, when it was omitted for the first time. In the 1960 volume a chapter on industrial psychology is scheduled to be written by Dr. B. von H. Gilmer. The contributors to the first nine years have all been well selected. All of the men have records of achievement as well as status. The list of contributors for the nine years includes: Shartle (1950), Bel- lows (1951), Brown and Ghiselli (1952), Harrell (1953), Heron (1954), Wallace and Weitz (1955), Kendall (1956), Katzell (1957), and Ferguson (1958).

The impression of most reviewers is that the editors of the *Annual Review* have done a good job of furnishing material to make possible the breaking of barriers for making understanding of each other to each other more possible, and have competently met this need. The question is whether it has been so used by the specialists in the various fields of psychology and, more particularly, for our present purpose, has it been so used by industrial psychologists who have reviewed the field for the *Annual Review* since its existence? *To examine the extent to which this has happened is one of the purposes of this study.*

D. DIFFERENTIAL USE OF THE TWO MAJOR REFERENCE SOURCES

A study of the tabulated data concerning the references used in the nine years of the chapters on industrial psychology in the *Annual Review* reveals some very interesting facts concerning the nature of the reading reference sources that the contributors generally used. It suggests some interesting facts about the scope that, to some extent, are alarming as well as encouraging. All of the contributors do a fair amount of reading. Not all of them have taken the freedom they have been invited to use with any degree of literary license or imagination. In all about 105 periodicals are referred to. Very few of them are mentioned year after year by each of the reviewers. There is one significant exception to this generalization. There are two major journals which are used year after year by every contributor and

which almost monopolize the field of industrial psychology as far as the reviews are concerned. The two journals are the *Journal of Applied Psychology*, which rates first in eight of the nine years, and *Personnel Psychology*, which rates first in only one of the nine years. However, the two journals jointly, with few exceptions, come close to averaging 50 per cent of all of the references used every year by each contributor. Because these two journals play such an important rôle in being the source of material used by the contributors to the *Annual Review*, the data indicating how each contributor uses each journal and the two journals combined is presented in Table 1. Frequency for each journal and for the two combined is reported both in terms of number of mentions and also per cent.

TABLE 1
TWO MAJOR REFERENCE SOURCES, 1950-1959

Item	1950	1951	1952	1953	1954	1955	1956	1957	1958
<i>Journal of Applied Psychology</i>	47	25	41	40	35	42	35	40	28
Percentage of total	36.8	25.5	32.5	28.6	27.8	34.4	24.3	31.0	11.6
<i>Personnel Psychology</i>	29	27	25	18	9	27	31	19	25
Percentage of total	22.6	27.5	19.9	12.8	7.1	22.1	21.5	14.7	10.4
Percentage of two major journals to total	59.4	53.0	52.4	41.4	34.9	56.5	45.8	45.7	22.0

From the data presented in Table 1, the following facts about the use of the two journals in the nine years are indicated. The range is from 22 per cent in 1958 to 59.4 per cent in 1950. In four of the nine years the percentage is above 50 per cent. From high to low, the years above 50 per cent are: 1950—59.4 per cent; 1955—56.5 per cent; 1951—53.0 per cent; 1952—52.4 per cent. In three of the years the percentage was above 40 per cent: 1956—45.8 per cent; 1957—45.7 per cent; and 1953—41.4 per cent. In 1954 it was 34.9 per cent and the lowest, 22.0 per cent, was in 1958. In general the contributors who are somewhat more traditional show a more frequent use of the two major journals as compared to other references. For example, Ferguson (1958) has by far the lowest percentage of use for the two journals. There is no doubt about his deviating most in his selection and organization and presentation of material. This doesn't necessarily mean that his chapter is better; it is just different. Also, it doesn't mean that it is more imaginative, just more deviated. The second lowest, Heron (1954), is easily explained by the fact that the references he uses are influenced by his English background. On the other hand, the contributors to the first three years all rated above 50 per cent and an analysis of the nature of the content reveals that they are more traditional in their orientation for selection and emphasis.

The *Journal of Applied Psychology*, which leads in every year but one, ranges in ratio of total mentions from the low of 11.6 per cent in 1958 to the high of 36.8 per cent in 1950. The most marked deviation here is in 1958, which is more than 10 per cent lower than the next lowest. The rest of the range is relatively smaller, namely from 24.3 per cent in 1956 to 36.8 per cent in 1950, with the average close to 30 per cent.

The second most frequently used journal is *Personnel Psychology*. The range in ratio of total mentions ranges from 7.1 per cent to 27.5 per cent. With the two lowest years omitted, the range is from 12.8 per cent to 27.5 per cent. In one year it rates the highest in being the journal most frequently used by the contributor to the *Annual Review*, but the measure of greater frequency is small—27.5 per cent for *Personnel Psychology* as compared to 25.5 per cent for the *Journal of Applied Psychology*. The lowest usage is found in Dr. Heron's (1954) contribution. If his is a sampling of English usage, this journal doesn't have the prestige and use in England that the *Journal of Applied Psychology* has. The next most marked deviation is again in 1958, with 10.4 per cent. The *Journal of Applied Psychology* is a more reliable index of traditional use of materials than is *Personnel Psychology*. The two combined are a still better index.

If there is a soaking process in putting oneself in shape as far as reading background for working up the chapter in the *Annual Review* is concerned, it is restricted in nature to the two journals most frequently mentioned as the chief sources. That there are two journals available which are subject to such frequent use is encouraging, but that the reading background is dependent on these two references to such a large extent is not a good sign of being well enough read for the purpose. Before we generalize too much on the basis of how these two major journals are used, let us first look at the distribution of use of all other references.

E. RANGE OF DISTRIBUTION OF REFERENCES IN THE NINE YEARS OF THE *Annual Review*

In addition to the two major journals most frequently used, what other journals are used and how frequently by how many of the contributors? In tabulated form the distribution of frequency of use of the 20 most frequently used references is shown in Table 2. Here are reported the total number of references for each journal for each year; the total for all years; the spread in years or the number of times out of the possible nine years that the journal was referred to and the number of sources or different periodicals used.

From a study of the data presented in Table 2, the facts concerning the

TABLE 2
DISTRIBUTIVE USE OF THE 20 MOST FREQUENTLY USED REFERENCES

DISTRIBUTIVE USE OF THE 20 MOST FREQUENTLY USED REFERENCES											
	Number of total mentions										
	1950	1951	1952	1953	1954	1955	1956	1957	1958		
1. <i>J. Appl. Psych.</i>	333	9	47	25	41	40	35	42	35	40	28
2. <i>Person. Psych.</i>	210	9	29	27	25	18	9	27	31	19	25
3. <i>Personnel</i>	70	9	5	2	9	6	1	14	10	12	11
4. <i>Occupational Psych.</i>	50	8	1	9	5	9	16	7	1	4	7
5. <i>Person. J.</i>	32	6	3		2	1	6	2			18
6. <i>Ed. & Psych. Meas.</i>	25	8	3		2	2	4	3	1	4	6
7. <i>Harvard Business Rev.</i>	18	4	3	1				1		11	5
8. <i>J. Marketing</i>	14	2								3	11
9. <i>Occupations</i>	13	5	1	4	4	2	2				6
10. <i>Pub. Person. Rev.</i>	12	3	1		5				1	1	7
11. <i>Human Rel.</i>	11	5	1			1	2	1	3	2	1
12. <i>Psych. Bull.</i>	10	6				3	3		10	4	
13. <i>J. Abn. & Soc. Psych.</i>	10	1		1			4		1	2	2
14. <i>Pers. Adm. & Ind. Rel.</i>	10	4						2	1		
15. <i>Ann. Rev. Psych.</i>	8	4				3			1		
16. <i>Fortune</i>	8	4				7	1				
17. <i>Bull. Ind. Psych. & Pers.</i>	8	2				1	2	1	1		
18. <i>Psych. Monog.</i>	7	5			2	1			5		
19. <i>Ind. Labor Rel. Rev.</i>	7	3			1	1					5
20. <i>Amer. Psychol.</i>	6	2				1					
Total reference to first 20	862	91	61	96	91	81	104	104	102	132	
Total reference to other periodicals	176	18	4	28	25	18	13	19	4	47	
Total reference to books	216	19	33	2	24	27	5	21	23	62	
Total references	1254	128	98	126	140	126	122	144	129	241	
No. periodical ref. sources	105	19	11	25	26	24	16	29	14	38	

major rôle played by the two journals already considered are obvious. Considered both in terms of frequency of mention only, or in frequency as well as spread, the two journals are in a class by themselves as compared to other sources of reference. The *Journal of Applied Psychology* received a total of 333 mentions and *Personnel Psychology* received 210 mentions. Both of them were referred to every year. *Personnel* is a poor third with 70 mentions, though it also is referred to in every year. Of all of the reference sources reported in Table 2, these three are the only ones mentioned in all nine years.

Two journals were referred to in eight of the nine years—*Occupational Psychology* and *Educational and Psychological Measurements*; two were referred to in six years—*Personnel Journal* and *Psychological Bulletin*; three were referred to in five years—*Occupations*, *Human Relations*, and *Psychological Monographs*; four were referred to in four years—*Harvard Business Review*, *Journal of Abnormal and Social Psychology*, *Annual Review of Psychology*, and *Fortune*; two were referred to in three years—*Public Personnel Review* and *Industrial Labor Relations Review*.

Occupational Psychology receives the fourth largest number of mentions and is referred to in eight of the nine years. The relatively high frequency score, however, is largely influenced by Heron's (1954) frequent use of the journal. His use of it is easily understandable and merely shows that an Englishman reads an English journal like Americans do not. *Personnel Journal*, mentioned in only six years, receives 32 mentions and ranks fifth in frequency of mention. *Educational and Psychological Measurements* has only 25 mentions, which ranks it sixth, though it is mentioned in eight years. Five journals were mentioned from 18 to 11 times; three were mentioned 10 times, and six were mentioned from 8 to 6 times. The rest of the mentions (not shown in Table 2) were five and below. The first 20 most frequent sources of reference range from 6 to 333 mentions. The *Journal of Applied Psychology* receives over 100 mentions more than the second most frequently used journal, *Personnel Psychology*, which in turn receives 140 more mentions than the third most frequently used journal, *Personnel*.

There are 10 instances in which the rank order shifts some when spread as well as frequency is considered. Information about how all of the 20 most frequently used references rank in terms of frequency only or with frequency as well as spread, can easily enough be obtained from the data presented in Table 2.

The *American Psychologist* ranks 20th on the basis of frequency of mentions and 19th if spread as well as frequency is considered. It is one publication that every psychologist receives and yet it is referred to in only two years—once by Heron (1954) and five times by Ferguson (1958). It

is interesting to note that though it is available to every psychologist, more so than every other journal, it ranks on the bottom of the first 20 references. This would indicate that apparently the *American Psychologist* hasn't had much in it in the last nine years which is of special interest to the contributors of the chapters on industrial psychology and is not considered a reference source for such information.

The *Psychological Bulletin* is referred to a total of 10 times—once by Bellows (1951), once by Harrell (1953), twice by Heron (1954), once by Wallace and Weitz (1955), three times by Kendall (1956), and twice by Katzell (1957). On rank order of frequency of mention it would be tied for 12th place with the *Journal of Abnormal and Social Psychology* and *Personnel Administration and Industrial Relations* though it was mentioned in six years while they were mentioned in four and one year respectively.

The *Psychological Review* is referred to once by Shartle (1950). The *Journal of Experimental Psychology* is referred to once by Harrell (1953). The *Journal of General Psychology* is referred to once by Bellows (1951). The *Journal of Psychology* is referred to five times by Brown and Ghiselli (1952).

Since social psychology is quite often given as a frame of reference for considering applications in industrial psychology, one would expect frequent use of social psychology sources to indicate this. Such is not the case. Of the journals in this area, the one most frequently referred to is the *Journal of Abnormal and Social Psychology*, which is referred to three times by Heron (1954), twice by Kendall (1956), four times by Katzell (1957), and once by Ferguson (1958). Other journals from this area rated significantly lower. For example, the *Journal of Social Psychology* is referred to only once by Shartle (1950) and twice by Heron (1954). The frequency of use of such related journals as the *American Journal of Sociology*, the *Journal of Social Issues*, *Sociometry*, and the *American Social Review* is low and not frequent enough to warrant either tabulation or recording of mention.

Clinical psychology in industry, considered in terms of emphasis on personality approaches and the use of clinical tools and techniques, has been increasingly emphasized in recent years. To what extent do any of the contributors show application of, familiarity with, or interest in clinical psychology in terms of references used? The *Journal of Clinical Psychology* was referred to only once, by Harrell (1953). The *Journal of Consulting Psychology* was referred to two times by Kendall (1956), and once by Ferguson (1958). Heron (1954) has a paragraph concerning the clinical method and the reference is to an article in the *Harvard Business Review* and the case study approach used at Harvard. Certainly this is no indication

that clinical psychology is influencing very many contributors to any significant degree.

Statistics is a tool that is often mentioned as a necessity for the industrial psychologist. References to sources of this nature include *Psychometrika*, which was referred to twice by Brown and Ghiselli (1952) and once by Heron (1954). *Biometrika* was referred to twice by Heron (1954). The *Journal of the American Statistical Association* was referred to once by Brown and Ghiselli (1952) and three times by Ferguson (1958). The *Journal of the Royal Statistical Association* was referred to once by Heron (1954). Again, familiarity with or frequency of use of statistics in the contributions is indicated in an extremely limited manner.

Human engineering receives fairly frequent mention in three of the early volumes. There are 13 references to human engineering by Shartle (1950), 17 by Brown and Ghiselli (1952), and 18 by Harrell (1953). It rates a chapter for itself for the first time in 1958. *Personnel Administration and Industrial Relations*, *Personnel Administration*, *Personnel Guidance Journal*, *Personnel Management*, *Personnel Research Bulletin*, and *Public Personnel Review* are mentioned by only a few of the contributors with generally low frequency.

Infrequent mentions ranging from one to five times by relatively few contributors are made to such periodicals as *Advertising Management*, *Public Opinion Quarterly*, *Public Administration Review*, *American Business*, *Management Record*, *Supervisory Management*, *Sales Management*, *Dunn's Review of Modern Industry*, *Journal of Economics*, *Operations Research*, and *Plant Administration*.

In the nine years quite a number of books on psychology and management were referred to. The variations from year to year are large. There is very little general knowledge that can be obtained by analyzing the number of book references used by the various contributors. To begin with, more often than not, the contributor for the year depends on the number of books that have come to his attention in the field or in related fields in the year. For example, Bellows (1951) referred to more than 20 chapters in the *Handbook of Applied Psychology* (6). Since one book is listed in the references more than 20 times, it means that he used relatively few books besides the *Handbook* since he had a total of 33 book references. Ferguson (1958) used the most books, as well as had the most references to periodicals. He has 62 book references. Considering the fact that his total contribution was 23 pages, of which 10 were reference pages, it can hardly be said that very much telling use of the references was made, except to have a long list. The most telling use of books referred to is made by Heron (1954). He

includes references to a book on work measurement by Abruzzi (1), a book on personality by Eysenck (4), and even an older volume on organization of character by Hartshorne, May, and Shuttlesworth (8). With relatively few exceptions, most of the books referred to were published in the year preceding the contribution. Rarely is an older volume referred to. Hop-pock's (10) *Job Satisfaction* is one of the older volumes referred to by Bellows (1951). The books selected, in general, are of relatively recent origin and rarely are they selected for historical perspective.

When factors like the scope of sensibility, precision of references, and depth of reasoning were rated by four judges, the 1954 contribution ranked first and the 1957 contribution ranked second. There was a high percentage of agreement, practically 90 per cent, in judging the 1954 contribution as best and 1957 as second best, with 1958 judged as being the lowest and the first three years as being relatively low. Those who rated relatively high had more organized perspective for selecting the content, making effective use of it with a purpose and objective in mind, and they manifested a more conscious awareness of what went before, what current progress was taking place, and the direction and meaning indicated by the studies reviewed. Heron (1954), who rated highest, used 24 different sources. The number of different sources used each year can be obtained from the data presented in Table 2. The range is from 11 in 1951 to 38 in 1958, with a median of 24. Aside from the 1958 contribution, which has already been described as merely having a long list without effective use of the list, the largest number of sources was used by Kendall (1956). A content review and analysis of his chapter indicates a breadth of consideration that is superior to most of the contributors. There is one characteristic he has which is more promising for the field. He has a familiarity with academic and industrial psychology and an interest in what they have to contribute to men in industry, as well as in what other related areas of knowledge have to contribute.

F. SUMMARY AND DISCUSSION

The availability of the *Annual Review* makes it possible for researchers in the field, as well as for students, to obtain, in a relatively fast manner, an orientation to the field considered in terms of reading background and orientation. To know what has gone before and to keep up with current progress is important for research workers and is even more important for the contributors to the *Annual Review* who have been invited to cover the material and also to evaluate and interpret the material so that ideas as well as facts concerning the field are made available. To investigate how well the contributors in industrial psychology have done in this respect is the

chief purpose of this study. For obtaining the facts needed, all of the references in the industrial psychology chapters for the nine years have been tabulated. Only the more general findings are reported and discussed in the summary.

1. The most outstanding fact found in studying the references for the nine years is that there are two major journal reference sources which are used year after year by every contributor. By far the most frequently used reference source is the *Journal of Applied Psychology*. In the first eight years it averages 30 per cent of the total references. The second most frequently used reference source is *Personnel Psychology*. It receives 210 mentions and averages close to 20 per cent of the total references. The two journals carry almost 50 per cent of the total references. The third most frequently used journal, *Personnel*, receives 140 mentions less than the second most frequently mentioned with 70 mentions. These three journals are the only ones that are used in each of the nine years. How frequently the other reference sources are used is presented in detail in Table 2.

2. The hope expressed by Kendler (13), that the *Annual Review* would add to the cohesiveness of an expansive field has, as far as the contributors to the industrial psychology chapters are concerned, not been fulfilled. Rarely do any of the contributors refer to articles considered in other chapters in the *Annual Review*, in spite of the fact that industrial psychology is considered the kind of field that needs to pick up content material from other areas. The consequence is that the availability of the various chapters, as far as industrial psychology is concerned, shows no sign of breaking the barriers between the various specialties.

Do the contributors to the industrial psychology chapters in the *Annual Review* get soaked in the literature before they cover the materials for interpretation and evaluation? More often than not it can be said that the tendency on the part of most of the contributors is not to be sufficiently soaked either in historical perspective or in related contemporary knowledge available in journals other than psychological. Being soaked in the literature does not happen to be a modern fashion with many of the writers of texts, as well as contributors to the *Annual Review*. For example, a recent text in psychology (18) emphasizes the importance of learning in its preface and has quite a number of chapters concerned with the topic. Yet E. L. Thorndike's name isn't even shown in the index! Generally speaking, all too often in contemporary psychological literature there is no evidence of familiarity with the field and related knowledge historically considered, as if what is currently being written is unrelated to what has gone before. Some of the contributions give the impression that the method of Paracelsus is with us again

in the field of psychology, namely that all knowledge that existed before the writers came on the scene is irrelevant and insignificant and that knowledge begins with their work. What Paracelsus did to show his contempt for what went before was to burn the books. What some psychologists do is neglect the reading of them.

Some specific suggestions for improvement that contributors to industrial psychology chapters can profit from, as implied in the findings, are:

1. Read chapters in the *Annual Review* that have contributions relevant for industrial psychology, as well as merely the chapters in industrial psychology.

2. Use relevant reference sources other than periodicals and books that are readily and easily available in psychological libraries and seminars.

3. Read books in literature and related fields such as economics, sociology, political science, and cultural anthropology in search for relevant source material.

4. Have in mind the need for a broader frame of reference, which includes a historical perspective to give the write-up the continuity of meaning which generally does not sufficiently characterize a number of the contributions in the nine years.

5. Read widely and be more selective and make more resourceful use of more relevant and significant sources rather than just either not reading or merely increasing the number of references without resourcefully using them.

6. Get soaked in literature in optimum terms, not maximum or minimum, but in terms of objectives in line with purposes and goals of the *Annual Review*.

7. Consciously attempt to integrate and correlate rather than merely digest and assimilate the facts and ideas of the sources used.

These are some of the specific suggestions that come to mind as guides for improvement. They obviously are not equally applicable to each contribution, but in a general way they do describe some of the needs that often appear and are implied in considering the contributions as a whole. The answer, of course, is not in favoring the more recent against the ancient, or the new against the old, but in being on the lookout for the relevant and significant.

About half of the contributors are relatively more traditional in their way of selective inclusiveness and the style of exclusiveness that goes with it. As a result they use source materials that are too confined to the boundary zones of psychological literature, or too easily available in psychological literature. Some of the contributors who deviate most, however, do not

always liberate their intelligence more effectively. They merely get distracted on an issue too important to them.

In general terms, what is needed even more than improvements indicated by the foregoing specific suggestions is a very general one that Muller (15) describes in "Exercises in Incongruity." He quotes Emerson as saying that it is a good thing now and then to take a look at the landscape between one's legs, and quotes Kenneth Burke as favoring this type of technique of analysis which he calls "perspective by incongruity." He argues that the evils of specialization cannot be eliminated by synthesis, because to harmonize and unify implies getting a perspective on perspectives. To do this he says the first move is to "escape the tyranny of any one scheme of thought by emphasizing the only thing that the whole history of philosophy, religion, and science makes plain: that a perspective is a perspective, not a transcendent vision of the whole, and that the drama of human thought is always presented from the restricted point of view of an actor in it, not of the omniscient author." The great revolutionary thinkers, like the great comic dramatists, he contends, are those who have most violently wrenched traditional associations, introduced perspective by incongruity on a large scale. Fascism is nothing but the last refuge of capitalism, religion is nothing but a device for keeping people quiet while skinning them, are illustrations given by Muller of the kind of generalizing that are in need of perspective by incongruity. There is occasion enough today for viewing with alarm, he informs us, but he reminds us it is we ourselves who create the occasion and cultivate the habit of alarm. Hence, he advocates "planned incongruity" in criticism as a practical social way of coming to feel more at home in the world we have to live in. This is a lesson that industrial psychologists can well apply in redirecting their energies as far as publications are concerned if the goal is to keep the field alive for growth in a healthy manner. The need is not for more generalities that still glitter, but for more genuine concern with more real problems which have significance for the present and for the future.

REFERENCES

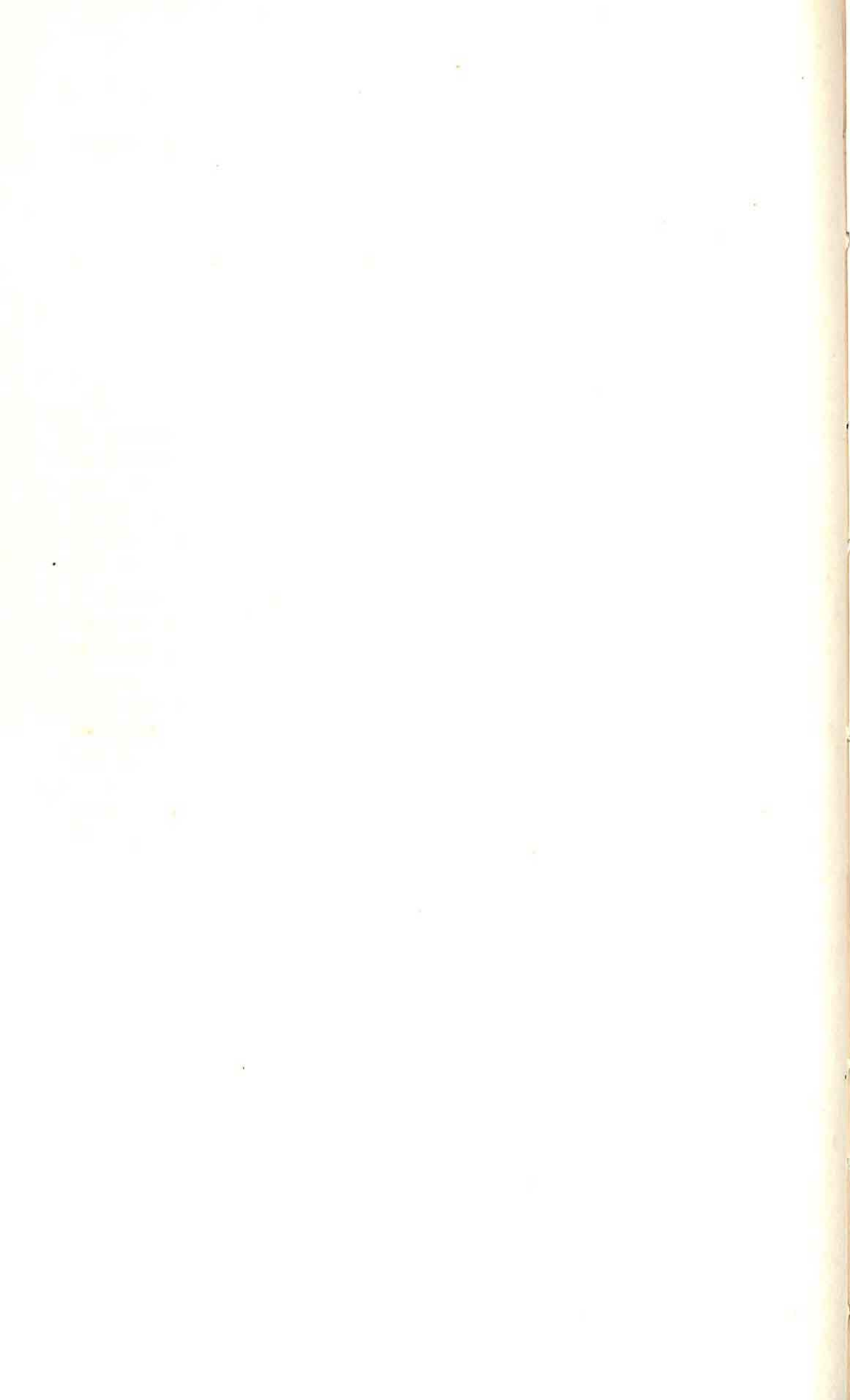
1. ABRUZZI, A. Work Measurement. New York: Columbia Univ. Press, 1952.
2. BELLOWS, R. M. Industrial psychology. *Annu. Rev. Psychol.*, 1951, **2**, 173-192.
3. BROWN, C. W., & GHISELLI, E. E. Industrial psychology. *Annu. Rev. Psychol.*, 1952, **3**, 205-232.
4. EYSENCK, H. J. The Scientific Study of Personality. London: Routledge and Kegan Paul, 1952.
5. FERGUSON, L. W. Industrial psychology. *Annu. Rev. Psychol.*, 1958, **9**, 243-266.
6. FRYER, D. H., & HENRY, E. R. Handbook of Applied Psychology. New York: Rinehart, 1950.

7. HARRELL, T. W. Industrial psychology. *Annu. Rev. Psychol.*, 1953, **4**, 215-238.
8. HARTSHORNE, H., MAY, M. A., & SHUTTLEWORTH, F. K. *Studies in the Organization of Character*. New York: Macmillan, 1930.
9. HERON, A. Industrial psychology. *Annu. Rev. Psychol.*, 1954, **5**, 203-228.
10. HOPPOCK, R. *Job Satisfaction*. New York: Harper, 1935.
11. KATZELL, R. A. Industrial psychology. *Annu. Rev. Psychol.*, 1957, **8**, 237-268.
12. KENDALL, W. E. Industrial psychology. *Annu. Rev. Psychol.*, 1956, **7**, 197-232.
13. KENDLER, H. H. Review of *Annu. Rev. Psychol.*, 1950. *Psychol. Bull.*, 1951, **48**, 159-161.
- 13A. MELTZER, H. Review of reviews in industrial psychology, 1950-1959. *Personnel Psychol.*, 1960, **13** (1).
14. MILLS, C. W. *The Sociological Imagination*. New York: Oxford Univ. Press, 1959.
15. MULLER, H. K. Exercises in incongruity. *Yale Review*, 1939, **29** (1), 97-112.
16. SHARTLE, C. L. Industrial psychology. *Annu. Rev. Psychol.*, 1950, **1**, 151-172.
17. WALLACE, S. R., & WEITZ, J. Industrial psychology. *Annu. Rev. Psychol.*, 1955, **6**, 217-250.
18. WICKENS, D. D., & MEYER, D. R. *Psychology*. New York: Dryden Press, 1955.

Human Relations Research Foundation

4510 Maryland Avenue

St. Louis 8, Missouri



LYSERGIC ACID DIETHYLAMIDE (*LSD-25*): XXX. THE QUESTIONNAIRE TECHNIQUE WITH NOTES ON ITS USE*

*Long Island Biological Association and the Community Hospital
at Glen Cove, New York*

H. A. ABRAMSON¹

A. INTRODUCTION

In response to numerous inquiries and to resolve misunderstandings, the questionnaire (3), which forms the basis of practically all of the publications on *LSD-25* by the writer and his colleagues, is herewith discussed. This questionnaire (Tables 1-2) was used in the form published here in the first paper and subsequent papers of this series, with one exception. It is also the test method of assessment in another series on *LSD* antagonists (2, 6). In the first paper published, only Part I of the questionnaire was analyzed statistically. Data were obtained, however, using Part II of the questionnaire, which was employed as the basis of a short mental rating test. Responses to Part II were later analyzed by Dr. Kornetsky (5).

The modified questionnaire, which we do not use, classified questions under physiological, perceptual, and cognitive (8). The writer feels that this type of classification robs the questionnaire of a certain degree of randomization that the original questionnaire possesses. The original questionnaire was composed by the writer and his assistant, Mrs. E. Teige, when the *LSD* Project was started about eight years ago. Mrs. Teige and the writer surveyed the literature on *LSD-25* and listed all the symptoms and signs reported. These symptoms and some of the signs, really the psychological signs, comprised the questionnaire now in use. As was expected, subsequent statistical studies revealed that certain of the questions were not so significant for the *LSD* reaction as were others. Another reason for using the questionnaire in its original form has to do with the nature of Part II, which as mentioned, can be used as a short mental rating test. Dependent on the experimental design, subjects may fill in Part I and Part II of the questionnaire, but they are also questioned by the observer. Additional

* Received in the Editorial Office on September 18, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This project has been aided by grants from the Josiah Macy, Jr., Foundation, New York, and the Sandoz Pharmaceuticals, Hanover, New Jersey.

TABLE 1

Half an hour after administration and at hourly intervals thereafter, the following direct questions are put to the subject. His rating is accepted, if possible, on a + to +++++ basis.

PART I QUESTION						
	½ Hr.	1½ Hrs.	2½ Hrs.	3½ Hrs.	4½ Hrs.	4½ Hrs.
1. Do you feel ill in any way?						
2. Are you nauseated?						
3. Have you a feeling of choking?						
4. Is salivation increased?						
5. Or decreased?						
6. Is your appetite increased?						
7. Or decreased?						
8. Do you have a "dry" taste in your mouth?						
9. Do you have a funny taste in your mouth?						
10. Is it a bitter taste?						
11. Are your lips numb?						
12. Or drawn back as if you were smiling?						
13. Does your head ache?						
14. Are things moving around you?						
15. Do you feel dizzy?						
16. Or unsteady?						
17. Is there difficulty in breathing?						
18. Do you pass more urine than usual?						
19. Are you aware of your heart beat?						
20. Is it faster than usual?						
21. Are you sweating?						
22. Are you hot?						
23. Or cold?						
24. Are your palms moist?						
25. Or dry?						
26. Or cold?						
27. Is your skin sensitive?						
28. Do you have funny feelings on your skin?						
29. Do your hands and feet feel peculiar?						
30. Do they feel heavy?						
31. Or light?						
32. Is there pressure in your ears?						
33. Is your hearing abnormal?						
34. Is it more acute than usual?						
35. Is your eyesight blurred?						
36. Do you have difficulty in focusing your vision?						
37. Do you see double?						
38. Are shapes & colors altered in any way?						
39. Does light bother you?						
40. Do things seem too close?						
41. Or too far away?						
42. Do you tremble inside?						
43. Do you feel weak?						
44. Or fatigued?						
45. Do you feel drowsy?						
46. Do you feel as if in a dream?						
47. Are you anxious?						

notes are made on the back of the cardboard test sheet (9.5" x 11.5") by the subjects themselves. Using Part I in this way, the questionnaire loses disadvantages often inherent in a strictly structured test situation and provides both the replicability of a questionnaire and the flexibility of a clinical assessment in addition.

TABLE 2

Half an hour after administration and at hourly intervals thereafter, the following qualitative ratings are made on the subject by the experimenter aided by the subject's comments.

PART II		
1. <u>Motor Behavior:</u>	2. <u>Control:</u>	3. <u>Consciousness:</u>
½ Hr.	½ Hr.	½ Hr.
1½ Hrs.	1½ Hrs.	1½ Hrs.
2½ Hrs.	2½ Hrs.	2½ Hrs.
3½ Hrs.	3½ Hrs.	3½ Hrs.
4½ Hrs.	4½ Hrs.	4½ Hrs.
Later	Later	Later
<u>Concentration:</u>	5. <u>Mood (Euphoria - Depression):</u>	6. <u>Attitude to Environment:</u>
½ Hr.	½ Hr.	½ Hr.
1½ Hrs.	1½ Hrs.	1½ Hrs.
2½ Hrs.	2½ Hrs.	2½ Hrs.
3½ Hrs.	3½ Hrs.	3½ Hrs.
4½ Hrs.	4½ Hrs.	4½ Hrs.
Later	Later	Later
7. <u>Orientation:</u>	8. <u>Memory:</u>	9. <u>Hallucinations:</u>
½ Hr.	½ Hr.	½ Hr.
1½ Hrs.	1½ Hrs.	1½ Hrs.
2½ Hrs.	2½ Hrs.	2½ Hrs.
3½ Hrs.	3½ Hrs.	3½ Hrs.
4½ Hrs.	4½ Hrs.	4½ Hrs.
Later	Later	Later

B. PROBLEMS

1. *Placebo Reactivity*

Only inadequate studies of psychotomimetic drugs can be made if poor communication exists between observer and test subject. It is recommended, therefore, that non-psychotic subjects be made aware of the content of the questionnaire before experiments are begun. Analysis of the data obtained on randomly selected, essentially non-psychotic subjects, showed the fre-

quency of placebo reactors to be high (3, 4). This demonstrated that the questionnaire technique screened out placebo positive subjects. It became apparent that one of the prerequisites for the study of the reaction of man to psychotomimetic drugs was to have essentially placebo-negative subjects. Indeed, even subjects who have been placebo negative may respond with symptoms in part characteristic of the *LSD* psychosis in times of psychological stress. Stress may exist either within or without the experimental situation. For this reason, repeated testing with placebos, or to save time and expense, with threshold doses of new compounds, is necessary.

2. Subject to Subject Variability

The questionnaire is useful in detecting the fact that in psychological studies large differences in reactivity to *LSD* and its congeners exist from subject to subject. These differences cannot be readily correlated with known variables. Various compounds may be given to each subject more than once, making two or three separate comparisons at separate dose levels for each drug, and thus have not only subject to subject data but also within subject studies (7). Although the effect of each drug will not be known as precisely as possible, these procedures by the use of the questionnaire provide us with adequate data. Subject to subject variability as well as within subject variability is illustrated in Table 3. For test Subject P.B.,

TABLE 3
THE EFFECT OF PSILOCYBIN ON TWO NON-PSYCHOTIC SUBJECTS

Subject	Experiment	Psilocybin (mg)	Questionnaire response (<i>n</i>)
P.B.	1	3	
	2	6	4
	3	6	2
	4	6	14
	5	6	13
D.V.G.		6	23
	1	6	
	2	6	16
	3	6	10
	4	8	10
			14

in a recent study of Psilocybin, the following data were obtained using Parts I and II of the questionnaire. The data disclose the wide variability in *n*, the number of responses, for Subject P.B. Compare these data of P.B. with studies on Subject D.V.G. with Psilocybin. Subject D.V.G. shows

much less within subject variability. These data are typical of the variability which may be disclosed using the questionnaire.

Our method of using the questionnaire enables us to compute a standard analysis of variance with three components: (a) Average difference between drugs for all subjects. (b) Subject to subject variation of psychotomimetic effects of the drugs. (c) An overall estimate of experimental error for each subject.

The real effectiveness of psychotomimetic differences from subject to subject can then be estimated.

3. *Subject Sophistication and Anxiety*

It must be recognized that a test group, on being used repeatedly, becomes somewhat sophisticated in their evaluation of symptom production. For example, in studying the effect of Prednisone on the LSD reaction, members of the test group could almost invariably detect whether or not Prednisone was being administered in advance by the occurrence of typical euphoria produced by Prednisone in the doses employed (2). This sophistication demonstrates, to a certain extent, the psychological threat invoked by the possibility of going psychotic in the test situation. In tests of tolerance and cross-tolerance, where subjects take large doses of psychotomimetic drugs for 10 days before the test situation itself, severe symptoms may occur. For example, following are verbatim remarks of Subjects R.B. and C.G. taken from the back of the questionnaire regarding the effects of Psilocybin during a 14-day tolerance experiment:

Subject R.B.: A considerable quantity of symptoms were experienced during the past week. The peak was reached on Monday and Tuesday of this week, having started on Saturday. The physical symptoms generally experienced were contraction of scalp, hollow feeling in pit of stomach, and unsteadiness. To walk purposefully was an effort. Monday I experienced great difficulty in concentrating on the conversations of others but was able to concentrate on work not involving others. A change in plans involving dinner preparations, in which I was to participate, completely "threw" me. I took some time to regain my composure. My temper was short throughout the day.

Tuesday I completely lost control of my temper at my child in the morning. Concentrated effort to avoid conflicts carried me through the rest of the day and since. Since Tuesday, the physical symptoms have been present, but the emotional symptoms were less severe. Yesterday and today I was able to carry on work involving others fairly well. Noontime symptoms are less severe than morning and evening symptoms. One must presume that some tolerance has been built up since symptoms have not increased with increasing dosages.

Subject C.G.: Saturday, Sunday, and Monday I had no particular reaction except wakefulness at night. Tuesday morning I had a two-and-a-half-hour reaction during which time I had nausea, dizziness, buzzing in the ears, light hands and feet, and considerable loss of time consciousness. Tuesday night I had the same feelings only they were worse, and I found I was unable to read a child's book with any coherency or sequence. The ideas weren't coming through.

Wednesday morning's reaction was worst of all. I sat at my desk, gazing out of the window for three hours, physically and mentally paralyzed. I felt as if I were anaesthetized. I couldn't answer questions or discuss anything. Wednesday night I split the dose in two parts, which was better. Thursday I split the dose in three parts in the morning and two parts in the evening, with only a slight effect each time. Friday the single dose had no effect at all.

Factors similar to the foregoing may structure the experiment on psychotomimetic drugs for the production of anxiety in the test subjects, independent of the drug itself. This situation is perhaps best illustrated by the fact that the queries of the questionnaire are very different indeed from the usual therapeutic trial where the expectation of the patient is to be helped, not to be threatened. For this reason, every effort is made to eliminate or dilute the anxiety, *which must be engendered by the test situation, no matter how dedicated or experienced the test subject may be.* A social milieu as free of stress as possible, with two or more subjects, is fairly anxiety free. Interaction by the members of the test group or their wives or husbands, who may be permitted to be present, diminishes anxiety because of the common bond produced by the test situation. It is of interest that during a television demonstration of the *LSD* psychosis by the writer the test subject wished his wife in the same room both before and during his television appearance (9).

To further prevent contamination of the test situation no visitors are permitted unless unanimous consent is obtained from members of the group. Test subjects are encouraged to verbalize recent experiences, which they feel may be anxiety producing, e.g., death of a close friend, occupational stress or nightmares. It has been found necessary to disclose occasionally to one member of the group information that partially destroys the "blind" or "double-blind" character of the experimental design. Thus, a test subject may be aware that he is really one of the first humans to have ever ingested a new congener of *LSD*. The test subject may ask, "Is this a new one?" This question should be answered with a noncommittal answer designed to reduce anxiety, but without disclosing precisely what is going on. A suitable answer might be, "Yes, but you are getting a low dose," or, "Yes,

but only an oxymethyl group has been added." It has been our experience that almost any answer will do, but rejection of the question must be avoided.

Ideally speaking, double-blind experiments may be desirable. However, a single-blind design produces much less anxiety. Tension arises in a subject unless he knows that someone responsible is consciously aware of his travail, incidental to the experiment. It is believed, therefore, that blind experiments are in general more suitable, although double-blind may be necessary at times.

It may be argued that continued experience with the same drug changes the reaction to the drug. However, as illustrated in Table 3, often the opposite may occur. For example, in recent experiments with Psilocybin, our subject P.B., for reasons unknown, during the first Psilocybin experience to 6 mg responded with only two positive responses on the questionnaire, whereas in the three subsequent Psilocybin experiments with 6 mg, the number of responses was 14, 13, and 23 respectively.

C. METHODS

1. *High Doses*

Too much reliance cannot be laid on Part I of the questionnaire with high doses. The greater the dose of a psychotomimetic drug, the more difficult it may be to communicate with the subject. In our present test group, used for the past five years, the threshold dose of *LSD* employed is about 25 mcgm. At 100 mcgm withdrawal symptoms with certain members of the group are great enough to produce symptoms so severe that problems arise. The second part of the questionnaire must be emphasized under these circumstances with the observer's evaluation often more important than the numerical value of the responses.

2. *Within Subject Studies*

Our first use of the questionnaire attempted to obtain a study of subject to subject variability with different doses. We have found that the questionnaire was especially useful for comparisons of psychotomimetic effects if studies were made within placebo negative subjects rather than between subjects. The method becomes sensitive if each subject receives all of the compounds at all of the dosage levels unless, as occasionally arises, a subject has a very low threshold and very severe reactions at higher doses. We believe that with this method variations within the response from subject to subject cancel out when response averages are compared. This leads to a gain in precision. Since the questions of the questionnaire are constructed so that a positive

response constitutes one response, by adding the number of responses a number, n , is obtained. It is obvious that, if the same drug is used in all experiments, different values of n may be compared. However, if *LSD* is to be compared with its congeners and derivatives, the significance of n only becomes important when the dose administered is introduced into the comparison. I have called the ratio of n divided by the dose in micrograms the Response Index, or

$$\text{Response Index} = R.I. = n/mcgm$$

For a given dose, therefore, the higher the value of *R.I.*, the greater the response of the drug. The use of the *R.I.* increases the sensitivity of the questionnaire and provides a suitable method of comparing psychotomimetic activity at threshold levels, or slightly above threshold levels, of various compounds related structurally. By obtaining the group average of the *R.I.*, a more representative value of psychotomimetic activity is established (1).

3. *Cross-Over Designs*

It is feasible to utilize a cross-over design in the study of various derivatives of *LSD*, varying the compounds within the group. The method has been especially effective in our present preliminary comparison of *LSD* and Psilocybin by administering both drugs to our test group double-blind. It may be mentioned that our test subjects cannot usually distinguish between *LSD* and Psilocybin taken orally, unless doses are near the threshold level.

4. *Grouping in Between Subject Comparisons*

We have found that age, sex, weight, occupation, psychological testing, pretest interviews, and screening by outside observers of pretest interviews did not provide adequate replicable groups. We are planning to explore this subject by comparing identical and nonidentical twins, both in our test group and also independently.

D. SUMMARY

The questionnaire used in assessing the effects of *LSD-25* and other psychotomimetic compounds in non-psychotic test subjects is discussed in detail. Its value in estimating placebo reactivity, subject to subject variability, subject sophistication and anxiety, within subject studies, and evaluation of a Response Index, is illustrated.

REFERENCES

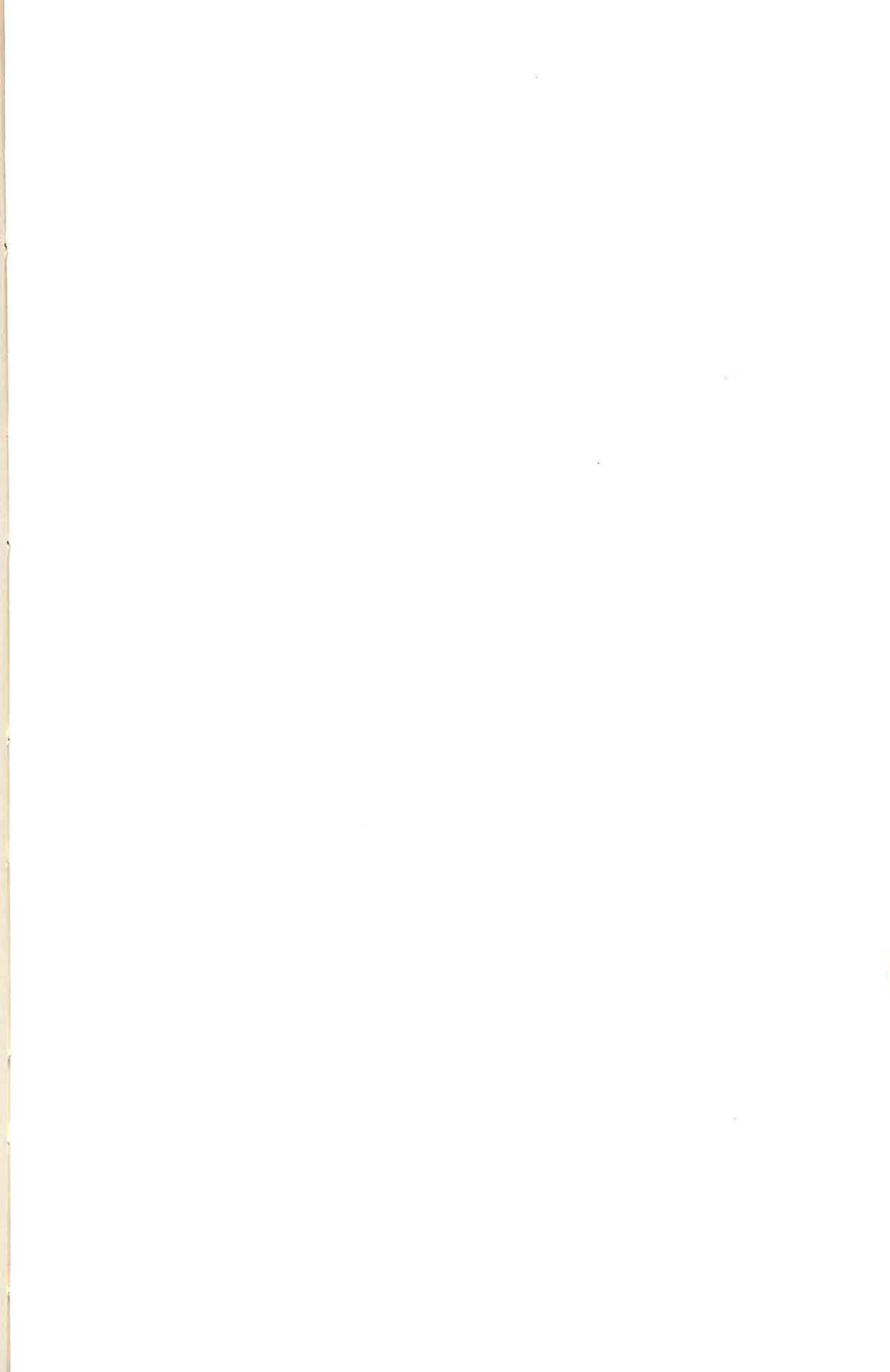
1. ABRAMSON, H. A. Lysergic Acid Diethylamide (LSD-25): XXIX. The Response Index as a measure of threshold activity of psychotropic drugs in man. *J. of Psychol.*, 1959, **48**, 65-78.
2. ———. Lysergic Acid Diethylamide (LSD-25) Antagonists: III. Modification of Syndrome by Prior Administration of Prednisone. *Arch. Gen. Psychiat.* (In press.)
3. ABRAMSON, H. A., JARVIK, M. E., KAUFMAN, M. R., KORNETSKY, C., LEVINE, A., & WAGNER, M. Lysergic Acid Diethylamide (LSD-25): I. Physiological and perceptual responses. *J. of Psychol.*, 1955, **39**, 3-60.
4. ABRAMSON, H. A., JARVIK, M. E., LEVINE, A., KAUFMAN, M. R., & HIRSCH, M. W. Lysergic Acid Diethylamide (LSD-25): XV. The effects produced by substitution of a tap water placebo. *J. of Psychol.*, 1955, **40**, 367-383.
5. ABRAMSON, H. A., KORNETSKY, C., JARVIK, M. E., KAUFMAN, M. R., & FERGUSON, M. W. Lysergic Acid Diethylamide (LSD-25): XI. Content analysis of clinical reactions. *J. of Psychol.*, 1955, **40**, 53-60.
6. ABRAMSON, H. A., SKLAROFSKY, B., BARON, M. O., & FREMONT-SMITH, N. Lysergic Acid Diethylamide (LSD-25) Antagonists: II. Development of tolerance in man to LSD-25 by prior administration of MLD-41 (1-Methyl-d-Lysergic Acid Diethylamide). *Arch. Neurol. & Psychiat.*, 1958, **79**, 201-207.
7. COCHRAN, W. G. Newer statistical methods. In *Quantitative Methods in Human Pharmacology and Therapeutics*. (Ed., D. R. Laurence.) New York: Pergamon Press, 1959.
8. JARVIK, M. E., ABRAMSON, H. A., & HIRSCH, M. W. *Comparative subjective effects of seven drugs including Lysergic Acid Diethylamide (LSD-25)*. *J. Abn. & Soc. Psychol.*, 1955, **51**, 657-662.
9. Mental Research Institute, Stanford University. The Fine Line. KPIX TV, San Francisco, California, August 18, 1959.

Biological Laboratory

Long Island Biological Association

Cold Spring Harbor, New York







INTENSIVE VITAMIN THERAPY IN MENTAL ILLNESS*

W. C. Kalash Laboratories, Pasadena, California

GEORGE WATSON AND W. D. CURRIER

A. PROBLEM

There is a mounting body of experimental evidence that some types of mental illness involve unsuspected nutritional deficiencies as well as psychological maladjustment. The literature on the relationships between nutrition and emotional disturbance has been reviewed recently by Bell (1). Experimental work in this area is of four principal types: (a) the creation of single vitamin deficiencies in humans with consequent psychological pathology; (b) the creation of a general state of malnutrition in human subjects with accompanying psychological disturbance (2, 4, 5); (c) the treatment with one or more vitamins of obviously undernourished individuals who also show signs of psychological malfunctioning; and (d) the treatment with vitamin and mineral therapy of psychologically diagnosed mentally ill individuals who exhibit no clinical signs of being undernourished. The study to be reported is in the last area, being the third in a continuing series of investigations of some of the relationships between mental illness and what appears to be nutritional need (7, 8).

B. METHOD

1. Procedure

Thirty volunteer unhospitalized emotionally disturbed subjects were selected for the study on the basis of having a history of emotional illness as well as having one or more abnormal scores on the Minnesota Multiphasic Personality Inventory (*MMPI*) (3, 6). Psychological case histories were taken in the initial interview, together with information concerning previous diagnoses and treatments when available. Every subject revealed a record of psychiatric or medical care for his condition.

Medical histories were obtained, and complete physical examinations and laboratory tests were given to all subjects. Periodic medical checks and laboratory tests were made as necessary.

The psychological progress of each subject was evaluated in terms of

* Received in the Editorial Office on September 18, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

changes in test scores on the Minnesota Multiphasic Personality Inventory, as well as by clinical psychological criteria.

Minnesota test results were interpreted in terms of a "total score" (7, 8), which was computed by summing the following scores: *Hs* plus .5*K*, *D*, *Hy*, *Pd* plus .4*K*, *Pa*, *Pt* plus 1*K*, *Sc* plus 1*K*, and *Ma* plus .2*K*.¹

Clinical evaluations were based upon the following criteria: (a) *Changes in attitudes and in behavior toward problems* (e.g., "I'm not so irritable toward my husband"); (b) *changes in real life situations* (e.g., "I quit my job and I'm going into business for myself"); (c) *renewal of faith in the future* (e.g., "I'm beginning to feel that everything will work out after all"); (d) *improvement in emotional stability* (e.g., "My father died yesterday—ordinarily I would have taken this very hard, but I'm not"); (e) *appraisal of self* (e.g., "I'm doing a lot better job than I've realized"); (f) *attention to physical appearance* (e.g., "I went out and bought a red dress yesterday for the first time in my life"); (g) *changes in specific symptoms of illness as reported in initial interview.*

All 30 subjects were first administered placebos. By having each subject act as his own control it was possible to make a direct comparison of the experimental treatment with the placebo effect. Subjects were transferred from placebos to experimental tablets without their knowledge when they exhibited no clinical signs of improvement in addition to having no significant changes in Minnesota Multiphasic Personality Inventory scores. Subjects who improved both clinically and on the *MMPI* while receiving placebos were continued on placebos until either they regressed or exhibited no further signs of progress.

Multiple amounts of the formula listed in Table 1 were given to every subject for all or for part of the experimental period. These were administered in tablet form, to be taken after meals. Subjects were given sufficient tablets at each interview for approximately one month's intake. Some subjects appeared to respond better to a formula differing from that given in Table 1, which did not contain methionine, vitamins A, B₁, B₂, or B₆ (see footnote, Table 1), while five subjects appeared to respond best to formulas reported in previous studies (7, 8)². The general problem of individual variability to nutritional replacement formulas will be discussed briefly below.

¹ *Hs* = hypochondriasis, *D* = depression, *Hy* = hysteria, *Pd* = psychopathic deviate, *Pa* = paranoia, *Pt* = psychasthenia, *Sc* = schizophrenia, *Ma* = hypomania, *K* = suppressor variable (a measure of distorting influence in test-taking attitude).

² Subjects 1, 4, 19, and 22 received daily four and one-half times the amounts of vitamins listed in Table 1, plus three times the minerals; Subjects 3, 7, 9, 12, 13, 14, 15, 17, 21, and 28 received daily three times the amounts of vitamins and minerals

TABLE 1
BASIC EXPERIMENTAL FORMULA*

	40,000 USP Units
Vitamin A	300 mg.
Vitamin C	0.5 mg.
Vitamin K	50 I.U.
Vitamin E	2 mg.
Vitamin B ₁	4 mg.
Vitamin B ₂	2 mg.
Vitamin B ₆	10 mg.
Vitamin B ₁₂	50 mg.
Intrinsic Factor Concentrate	320 mg.
Niacinamide	1.5 mg.
Folic Acid	108 mg.
Calcium Pantothenate	50 mg.
Lemon Bioflavonoid Complex	33 mg.
Inositol	400 mg.
Choline Bitartrate	400 mg.
Betaine Monohydrate	67 mg.
Methionine	500 mg.
Calcium	375 mg.
Phosphorus	10 mg.
Iron	0.3 mg.
Iodine	0.15 mg.
Copper	0.1 mg.
Nickel	0.1 mg.
Zinc	0.1 mg.
Molybdenum	0.25 mg.
Fluorine	0.8 mg.
Sodium	0.5 mg.
Sulfur	5 mg.
Manganese	5 mg.
Magnesium	0.75 mg.
Potassium	0.25 mg.
Cobalt	

* The following is a variation of the basic formula received by some subjects (see Footnote 2):

Vitamin B₁₂ with Intrinsic Factor Concentrate, 1 USP Units (oral); Vitamin C, 200 mg.; Folic Acid, 1 mg.; d-Panthenol, 100 mg.; Inositol, 120 mg.; Choline Bitartrate, 120 mg.; Niacinamide, 200 mg.; Vitamin E, 120 Int. Units; Vitamin K, 0.4 mg.; Betaine Monohydrate, 120 mg.; Lemon Bioflavonoid Complex, 200 mg.; Hesperidin, 100 mg.; Rutin, 50 mg.; Calcium, 200 mg.; Phosphorus, 150 mg.; Iron, 24 mg.; Iodine, 0.2 mg.; Copper, 0.1 mg.; Zinc, 0.08 mg.; Cobalt Chloride, 4 mg.; Sea Salt Trace Elements, 30 mg. (in three tablets).

The vitamin formulas used in this study were supplied by the Kalash Vitamin Corporation, Pasadena 4, California.

listed in Table 1; Subjects 18, 23, 30 received the formula listed in the footnote to Table 1; Subjects 5 and 11 received daily twice the quantities listed in the Table 1 footnote; Subjects 2, 6, 8, 16, 20, and 29 received daily three times the amounts listed in the footnote, Table 1; Subjects 10, 24, 26, and 27 responded best to the formula employed in (7), while Subject 25 responded best to the formula employed in (8).

C. RESULTS

Table 2 lists the Minnesota Multiphasic Personality Inventory scores for the beginning and the end of the placebo period. The average length of the placebo trial was 2.46 months (range 1-14 months, median one month).

Of 30 placebo subjects, seven improved, six got worse, and 17 showed no change, judged on a clinical basis. The clinical improvements for the placebo group are not statistically significant. Evaluated in terms of improvement on the Minnesota Multiphasic Personality Inventory, the average reduction in total test score for 30 subjects was 4.40 points. This difference is not statistically significant.

Table 3 lists the *MMPI* scores for the end of the placebo period and the end of the first experimental period. The first experimental period represents an average time interval of 1.76 months (range 1-15 months, median one month). This average period of time was determined by matching the lengths of individual placebo trials with as nearly equal periods of experimental trials as possible.

Clinical evaluations at the end of the first experimental period indicated 22 subjects improved, two worse, and six unchanged. This is significant well beyond the .01 level of confidence. Judged in terms of the Minnesota Multiphasic Personality Inventory, the average reduction in total score for the first experimental period was 17.13 points for 30 subjects. This improvement is also significant well beyond the .01 level of confidence.

Table 4 lists the *MMPI* scores for the end of the placebo period and the end of the total experimental period. The average length of the total experimental period was ten months (range 1-30 months, median 8 months). Variations in the length of the total experimental period for individual subjects were due to the fact that some subjects entered the program considerably later than others, while some received placebos for longer intervals than others, reducing the lengths of their experimental periods.

Clinical evaluations at the completion of the experiment indicated 24 subjects improved, one worse, and five unchanged. This is significant well beyond the .01 level of confidence. Evaluated in terms of improvement on the Minnesota Multiphasic Personality Inventory, the average reduction in total test score was 26.77 points for 30 subjects, which again is significant well beyond the .01 level of confidence.

Comparing the effect of the total experimental treatment directly with the placebo effect (net change in total score on the Minnesota Multiphasic Personality Inventory on experimentals minus net change in *MMPI* total

TABLE 2
 MINNESOTA MULTIPHASIC PERSONALITY INVENTORY RESPONSE TO PLACEBO TRIAL*

Subject	Age	Sex	F	K	HS + .5K	D	Hy	Pd + .4K	Pa	Pt + 1K	Sc + 1K	Ma + .2K	Total score	Months on placebos
1	45	F	15	10	11	41	20	32	13	48	51	22	238	1
			19	9	16	40	26	29	14	46	54	20	245	
2	32	M	5	12	24	28	29	21	7	32	25	12	178	1
			4	19	25	28	29	26	8	32	32	14	194	
3	31	F	13	16	15	38	32	36	11	40	38	23	233	1
			11	15	12	34	28	36	13	40	40	24	227	
4	34	F	9	12	27	34	34	25	10	41	21	16	208	2
			6	12	28	34	30	23	10	38	29	19	211	
5	20	F	8	12	18	28	30	32	17	34	38	26	223	1
			10	10	20	23	28	33	17	45	43	32	241	
6	60	F	6	18	34	50	44	26	17	45	46	21	283	1
			6	19	34	45	39	28	15	47	43	20	271	
7	29	M	14	23	30	35	40	30	13	36	46	23	253	1
			16	17	29	39	35	32	17	35	47	19	253	
8	56	M	10	10	14	40	22	22	11	39	34	17	199	1
			8	9	14	38	20	19	12	38	31	19	191	
9	49	M	4	20	22	22	31	27	14	35	35	19	205	1
			5	19	19	28	31	30	13	37	36	24	218	
10	28	F	6	15	18	28	28	26	14	34	27	17	192	1
			4	19	19	24	22	28	13	32	30	22	190	
11	50	F	4	14	28	28	33	27	13	36	36	24	225	1
			4	13	25	29	36	32	14	40	33	28	237	
12	38	F	16	7	23	30	29	28	17	33	43	21	224	2
			15	10	26	29	35	33	12	36	44	22	237	
13	45	F	12	6	12	33	23	26	20	36	41	26	217	1
			15	16	20	29	31	29	13	36	40	29	227	
14	55	F	10	12	20	33	32	25	18	42	44	23	237	1
			7	16	22	40	32	29	17	37	40	20	237	
15	37	M	4	11	24	28	33	25	10	40	31	18	209	2
			1	11	22	27	31	25	13	42	29	14	203	
16	59	F	7	23	34	34	43	21	16	42	43	13	246	1
			8	24	36	31	43	23	14	41	46	12	246	
17	29	F	7	10	14	16	19	23	12	31	30	31	176	3
			4	11	10	18	21	26	16	29	24	24	168	
18	56	F	3	17	32	35	37	16	12	42	31	13	218	4
			3	21	30	37	40	22	13	49	36	20	247	
19	50	M	5	11	13	39	15	17	11	35	25	13	168	1
			6	12	13	44	24	20	12	42	29	9	193	
20	59	F	9	15	28	39	39	29	14	36	32	15	232	1
			4	17	28	34	37	24	11	39	30	16	219	
21	35	M	9	17	11	21	25	33	15	33	37	26	201	2
			6	13	10	30	27	35	14	38	38	27	219	
22	47	F	6	19	21	26	27	25	13	31	41	32	216	1
			5	15	23	36	31	15	15	38	38	19	215	
23	19	F	13	5	19	29	29	25	14	43	46	28	233	14
			7	12	13	28	24	26	15	43	33	19	201	

* The first listed total score for each subject is for his initial test; the second is his total score at the end of the placebo period.

TABLE 2 (Continued)
MINNESOTA MULTIPHASIC PERSONALITY INVENTORY RESPONSE TO PLACEBO TRIAL

Subject	Age	Sex	F	K	Hs + .5K	D	Hy	Pd + .4K	Pa	Pt + 1K	Sc + 1K	Ma + .2K	Total score	Months on placebos
24	58	M	6	10	29	40	34	23	6	38	28	13	211	2
25	36	F	4	9	24	41	29	17	10	35	20	8	184	
			7	12	23	28	33	32	13	52	42	28	251	5
			5	9	16	26	26	31	12	37	28	28	204	
26	40	F	2	12	16	31	23	22	11	37	34	18	192	2
			3	13	13	24	18	18	7	32	28	22	162	
27	32	F	8	10	28	31	34	27	12	38	38	22	230	3
			3	15	23	23	27	21	10	39	41	19	203	
28	70	F	15	14	26	37	41	26	18	46	45	20	259	1
			11	12	21	34	38	23	17	42	37	23	235	
29	52	F	2	13	22	28	30	27	15	34	38	21	215	5
			1	22	24	16	30	24	12	30	28	21	185	
30	57	F	6	19	24	33	33	23	9	33	28	16	199	11
			2	18	19	26	29	21	9	32	24	16	176	

TABLE 3
MINNESOTA MULTIPHASIC PERSONALITY INVENTORY RESPONSE TO FIRST EXPERIMENTAL TRIAL*

Subject	Age	Sex	F	K	Hs + .5K	D	Hy	Pd + .4K	Pa	Pt + 1K	Sc + 1K	Ma + .2K	Total score	First experimental period (months)
1	45	F	19	9	16	40	26	29	14	46	54	20	245	1
2	32	M	13	12	12	35	22	26	10	36	32	16	189	
			4	19	25	28	29	26	8	32	32	14	194	1
3	31	F	5	20	17	22	23	26	8	31	26	13	166	
			11	15	12	34	28	36	13	40	40	24	227	1
4	34	F	9	13	11	29	28	31	14	37	33	23	206	
			6	12	28	34	30	23	10	38	29	19	211	2
5	20	F	3	14	22	31	28	25	13	35	23	22	199	
			10	10	20	23	28	33	17	45	43	32	241	1
6	60	F	9	6	15	29	25	29	14	40	39	28	219	
			6	19	34	45	39	28	15	47	43	20	271	1
7	29	M	6	18	32	41	39	27	13	44	42	20	258	
			16	17	29	39	35	32	17	35	47	19	253	1
8	56	M	10	17	24	33	32	27	12	31	43	23	225	
			8	9	14	38	20	19	12	38	31	19	191	1
			8	11	16	36	21	18	4	39	31	20	185	

* The first listed total score for each subject is for his initial test at the start of the experimental period; the second is his score at the end of a period most nearly equalling the length of his placebo trial.

TABLE 3 (Continued)
MINNESOTA MULTIPHASIC PERSONALITY INVENTORY RESPONSE TO FIRST
EXPERIMENTAL TRIAL

Subject	Age	Sex	F	K	Hs + .5K	D	Hy	Pd + .4K	Pa	Pt + 1K	Sc + 1K	Ma + .2K	Total score	First experimental period (months)
9	49	M	5	19	19	28	31	30	13	37	36	24	218	1
			4	26	21	23	29	24	11	30	42	19	199	
10	28	F	4	19	19	24	22	28	13	32	30	22	190	1
			1	17	18	22	23	22	9	30	30	20	174	
11	50	F	4	13	25	29	36	32	14	40	30	28	237	1
			6	16	32	26	35	30	12	39	44	29	247	
12	38	F	15	10	26	29	35	33	12	36	44	22	237	2
			15	8	27	26	28	30	14	33	43	26	227	
13	45	F	15	16	20	29	31	29	13	36	40	29	227	1
			6	16	14	25	14	26	12	27	24	20	162	
14	55	F	10	12	20	33	32	25	18	42	44	23	237	1
			8	12	18	36	27	28	15	43	39	21	227	
15	37	M	1	11	22	27	31	25	13	42	29	14	203	2
			3	10	22	25	31	20	13	40	31	21	203	
16	59	F	8	24	36	31	43	23	14	41	41	15	246	1
			8	24	35	32	42	26	14	41	41	15	246	
17	29	F	4	11	10	18	21	26	16	29	24	24	168	1
			10	10	11	19	19	23	12	32	34	23	173	
18	56	F	3	21	30	37	40	22	13	49	36	20	247	5
			2	20	25	34	31	17	12	42	29	13	203	
19	50	M	6	12	13	44	24	20	12	42	29	9	193	1
			3	10	9	31	14	16	9	32	21	15	147	
20	59	F	4	17	28	34	37	24	11	39	30	16	219	5
			3	16	18	31	28	28	11	33	22	14	185	
21	35	M	6	13	10	30	27	35	14	38	38	27	219	1
			4	18	9	20	20	26	10	26	29	23	163	
22	47	F	5	15	23	36	31	15	15	38	38	19	215	1
			3	16	13	28	20	20	12	32	27	18	170	
23	19	F	7	12	13	28	24	26	15	43	33	19	201	3
			3	14	12	29	25	24	13	38	31	17	189	
24	58	M	4	9	24	41	29	17	10	35	20	8	184	3
			5	15	26	41	34	22	11	38	29	10	211	
25	36	F	5	9	16	26	26	31	12	37	28	28	204	3
			6	12	11	27	22	28	11	39	33	28	199	
26	40	F	3	13	13	24	18	18	7	32	28	22	162	2
			3	13	14	24	21	21	7	29	27	17	160	
27	32	F	3	15	23	23	27	21	10	39	41	19	203	2
			9	12	21	25	20	22	8	36	41	20	193	
28	70	F	11	12	21	34	38	23	17	42	37	23	235	1
			9	12	21	35	35	28	16	47	39	21	242	
29	52	F	1	22	24	16	30	24	12	30	28	21	185	4
			0	20	24	14	33	24	14	28	26	19	182	
30	57	F	2	18	19	26	29	21	9	32	24	16	176	2
			4	16	19	29	32	23	10	28	19	13	173	

TABLE 4
MINNESOTA MULTIPHASIC PERSONALITY INVENTORY RESPONSE TO TOTAL
EXPERIMENTAL TRIAL*

Subject	Age	Sex	F	K	Hs + .5K	D	Hy	Pd + .4K	Pa	Pt + 1K	Sc + 1K	Ma + .2K	Total score	Total months experimental period
1	45	F	19	9	16	40	26	29	14	46	54	20	245	22
			9	17	10	30	16	24	12	31	36	13	172	
2	32	M	4	19	25	28	29	26	8	32	32	14	194	22
			4	21	14	21	20	24	8	22	25	16	150	
3	31	F	11	15	12	34	28	36	13	40	40	24	227	24
			8	15	11	31	25	28	9	34	36	20	194	
4	34	F	6	12	28	34	30	23	10	38	29	19	211	10
			2	14	23	26	25	22	7	35	20	21	179	
5	20	F	10	10	20	23	28	33	17	45	43	32	241	8
			6	8	13	27	25	28	15	39	27	31	205	
6	60	F	6	19	34	45	39	28	15	47	43	20	271	13
			5	21	32	42	37	27	13	38	35	18	242	
7	29	M	16	17	29	39	35	32	17	35	47	19	253	27
			12	13	18	27	21	20	8	29	33	21	177	
8	56	M	8	9	14	38	20	19	12	38	31	19	191	8
			10	10	16	36	26	17	5	36	27	21	184	
9	49	M	5	19	19	28	31	30	13	37	36	24	218	7
			1	26	13	18	24	22	9	27	27	18	158	
10	28	F	4	19	19	24	22	28	13	32	30	22	190	11
			3	18	16	23	25	26	11	30	27	21	179	
11	50	F	4	13	25	29	36	32	14	40	33	28	237	15
			3	19	24	24	35	21	10	34	27	25	200	
12	38	F	15	10	26	29	35	33	12	36	44	22	237	8
			11	9	18	31	24	32	11	31	37	22	206	
13	45	F	15	16	20	29	31	29	13	36	40	29	227	10
			1	18	12	21	19	23	15	31	28	21	170	
14	55	F	7	16	22	40	32	29	17	37	40	20	237	8
			4	14	21	40	34	25	16	43	33	18	230	
15	37	M	1	11	22	27	31	25	13	42	29	14	203	11
			4	14	24	25	35	25	13	40	30	16	208	
16	59	F	8	24	36	31	43	23	14	41	46	12	246	9
			6	24	32	28	37	25	14	39	40	17	232	
17	29	F	4	11	10	18	21	26	16	29	24	24	168	4
			2	7	7	12	16	23	12	26	28	24	148	
18	56	F	3	21	30	37	40	22	13	49	36	20	247	5
			2	20	25	34	31	17	12	42	29	13	203	
19	50	M	6	12	13	44	24	20	12	42	29	9	193	10
			2	14	9	26	17	19	7	23	18	19	138	
20	59	F	4	17	28	34	37	24	11	39	30	16	219	11
			3	18	24	35	34	29	11	38	27	12	210	
21	35	M	6	13	10	30	27	35	14	38	38	27	219	3
			4	17	10	20	23	27	8	30	30	20	168	
22	47	F	5	15	23	36	31	15	15	38	38	19	215	7
			3	24	17	19	22	22	9	26	28	20	163	
23	19	F	7	12	13	28	24	26	15	43	33	19	201	3
			3	14	12	29	25	24	13	38	31	17	189	

* The first listed total score for each subject is for his initial test at the start of the experimental period; the second is his score at the end of the experiment.

TABLE 4 (Continued)
 MINNESOTA MULTIPHASIC PERSONALITY INVENTORY RESPONSE TO TOTAL
 EXPERIMENTAL TRIAL

Subject	Age	Sex	F	K	Hs + .5K	D	Hy	Pd + .4K	Pa	Pt + 1K	Sc + 1K	Ma + .2K	Total score	Total months experimental period
24	58	M	4	9	24	41	29	17	10	35	20	8	184	8
			4	9	23	44	32	19	9	32	24	8	191	
			4	9	26	26	26	31	12	37	28	28	204	3
25	36	F	5	9	16	26	25	26	10	32	26	24	186	
			4	9	17	26	25	26	10	32	28	22	162	10
26	40	F	3	13	13	24	18	18	7	31	28	21	169	
			3	15	13	27	20	22	7	39	41	19	203	3
27	32	F	3	15	23	23	27	21	10	35	36	18	186	
			7	15	20	26	20	23	8	42	37	23	235	21
28	70	F	11	12	21	34	38	23	17	42	42	23	237	
			10	14	22	31	35	25	15	44	28	21	185	4
29	52	F	1	22	24	16	30	24	12	30	28	21	189	
			0	20	22	17	31	26	13	32	27	21	176	1
30	57	F	2	18	19	26	29	21	9	32	24	16	173	
			4	16	19	29	32	23	10	28	19	13		

score on placebos) the mean net difference for the experimental treatment was 22.37 points, a difference which is significant beyond the .01 level of confidence.

D. SELECTED CASE HISTORIES

Three case histories have been selected for summary because of the intrinsic interest of the psychopathologies represented. An earlier study (7) presents more detailed information on degrees and types of clinical response related to total score changes on *MMPI*.

1. Subject 7, Male, Age 29

He reported a history of progressively increasing illness over a 12-year period, beginning with auditory hallucinations when he was in high school, followed at the age of 25 by a period of amnesia and an acute psychotic episode for which he was hospitalized. He was given approximately 20 shock treatments, followed by tranquilizers and by analysis, two periods per week, for one and one-half years. He had been diagnosed as a paranoid schizophrenic.

In the initial interview he stated that his thinking as well as his behavior was "controlled by a machine in outer space." He said: "It's like tyranny. Sometimes the machine holds me down on the bed and won't let me get up." He also reported that there was a voice in his head that talked constantly.

"It talks about many things—intelligent, sometimes cruel." He further said: "The machine controls my voice and it also has a voice. This thing is one me (i.e., the machine)." He also reported continuous pain in the bowels, stomach, legs, and an "over-all throbbing in my head."

His initial total score on the Minnesota Multiphasic Personality Inventory was 253, with abnormal peaks on hypochondriasis, depression, hysteria, psychopathic deviation, psychasthenia, and schizophrenia (see Table 2). At the end of a one-month period on placebos his *MMPI* total score remained unchanged at 253, and he stated that he couldn't tell whether or not the tablets were helping him.

At the end of the first month after being switched to experimental tablets without his knowledge he again said "I can't say whether the tablets are helping," but his *MMPI* total score improved by 28 points (from 253 to 225). He also stated that "The consciousness of the machine is lessening" and "the voice has lessened in intensity." When questioned about the machine which controlled him from outer space, he now said "the machine is me."

At the end of the second month on the experimental tablets he repeated, "The machine is me." And then he said: "Looking back I didn't recognize it before. This thing has to be in my own head—it couldn't possibly be outside." He further stated: "The voice doesn't interrupt when I'm looking at TV now. I'm more on the road to recovery than I was before—more alert and attentive."

Fluctuations in his *MMPI* scores were noted when the formula he was taking was varied, as well as after periods of unusual stress, such as the Christmas holidays.³ Despite such setbacks, however, at the end of the first seven months he had gained 15 pounds, and seemed improved clinically. He was no longer hostile, aloof, or suspicious, and his sense of humor began to return. He stated: "I've felt better this month than I have for a long time. The pains all over have diminished. I have not been awake at night like I used to." He previously had reported that "the voice" talked to him for hours upon retiring, while now he was going to sleep within 15 minutes without this happening. His mother reported that while he had "some depressed days," he "was beginning to act like he did before becoming ill,"

³ Approximately six months of the 30-month experimental period was consumed in making variations in his formula in an attempt to obtain optimum response. These changes are too complex to be detailed, but in general, the addition of quantities of vitamins B₁, B₂, B₆, in excess of the proportions listed in Table 1 seemed to be undesirable. Further, vitamin A, as well as Inositol, Choline, Betaine, Methionine, appeared important. When they were eliminated, or sharply reduced in quantity, his clinical response as well as his *MMPI* scores were poorer (see Footnote 2).

and that he had voluntarily gone to the library for the first time in a couple of years and was studying physics in his leisure.

At the end of the experimental period his *MMPI* total score had improved by 76 points, from 253 to 177. He reported being able to see that his test scores were improving, stating "I am able to be much more objective on my answers." Although not fully recovered, he had made considerable improvement clinically.

2. *Subject 15, Male, Age 37*

He reported a history of severe anxiety and depression extending over a five-year period. He had received psychotherapy from an analyst for about one year, followed by treatment by an internist. At no time was his illness severe enough to require hospitalization or to cause him to be absent from his work, although he stated: "I don't know how I ever got by the years 1953-54. I still wonder how I get along." He reported being dizzy, seeing the "walls moving," being extremely fatigued with intermittent periods of excitement, and thinking constantly of dying.

His first total score on the Minnesota Multiphasic Personality Inventory was 209. After two months on placebos, it was 203, and he said: "I've felt worse this month—certainly no better." After being switched without his knowledge to experimental tablets for one month he seemed to be slightly improved, and his *MMPI* score dropped temporarily. Subsequent retests showed no improvement over his initial score. He continued to report periods of extreme nervousness and anxiety, which none of the variations in experimental formulas were successful in helping over a one-year period. At the end of the experiment he was judged clinically unimproved.

3. *Subject 22, Female, Age 47*

She reported a long history of emotional disturbance, including two psychotic episodes, the latest of which occurred several weeks before the beginning of her participation in the study here reported. She had been under general medical as well as psychiatric care for the past several years, and had been taking a variety of tranquilizers.

In the initial interview she reported that during acute periods of illness "I felt like I had two minds." She also said: "I went for six weeks like another person. What I look for when I am that other person is an ideal man." She blamed her husband for her condition, stating that lack of sympathy, understanding, and companionship caused her to look for other men. She also complained: "I had to give up all my family since my husband doesn't approve of them—seems like I don't have anyone."

Her husband reported in a separate interview that his wife during her more disturbed periods would take up with other men for "companionship." He stated that these friendships were not extra-marital "affairs." He also revealed that his wife had two sisters, one of whom would "wander away," picking up other men, while the other behaved similarly and had been confined to a mental hospital. He added that his wife was just like her sisters, and his behavior was not the cause of her illness.

Her initial total score on the Minnesota Multiphasic Personality Inventory was 216, the highest peak being hypomania, in addition to having an abnormal schizophrenia score. At the end of six weeks on placebos, she reported being slightly improved in some ways, but more depressed and apathetic. Her total score on the *MMPI* remained about the same (215), with the profile revealing more depression. At the end of one month after being switched to experimental tablets without her knowledge, she reported: "I am better, much better. I feel better than I've felt in seven or eight years. I never expected to feel this good again." Her score on the Minnesota test dropped 45 points to a total of 170 in the first month of treatment, and the entire profile was now normal.

At the end of the second month's treatment she volunteered: "I feel normal. I'm getting along fine with my husband. I have about as ideal a husband as I can find. I could look the whole world over and not find a better one than he is."

Re-tests on the *MMPI* during the experimental period averaged 170 points in total score (range 163-180), with her total at the end of the experiment being a normal 163. At that time, after eight months on the experimental regimen, she reported: "I'm fine, wonderful. My husband's fine. I can never understand how one person can interrupt a home life like I did. My father died yesterday; ordinarily I would have taken this very hard, but I'm not."

E. DISCUSSION

The results of the present study appear to be superior to those previously reported (7, 8), both in the number of subjects showing marked clinical improvements as well as in the amount of average reduction in total test score on the Minnesota Multiphasic Personality Inventory.

The apparent superiority of these results may be due to differences in the basic formula employed as compared to those used previously. These differences may be summarized as follows: vitamins A, E, C, pantothenic acid, niacinamide, cyanocobalamin, folic acid, choline, betaine, and inositol,

when employed in much larger quantities proportionally to thiamine, riboflavin, and pyridoxine, and with para-aminobenzoic acid and vitamin D omitted, appear to give more satisfactory results in most cases. However, some subjects respond best to the nutritional complex employed in (7), which has larger amounts of B₁, B₂, B₆, in addition to containing vitamins D and PABA, and much smaller amounts proportionally of A, E, C, pantothenic acid, B₁₂, folic acid, choline, betaine, and inositol. In addition, a formula in which vitamins A, B₁, B₂, B₆, PABA, and D are eliminated, as in the complex listed in the footnote to Table 1, is apparently markedly superior for some subjects. In this connection it is important to note that an incorrectly balanced nutritional complex may aggravate symptoms. In the study here reported the problem of variation in clinical response to a given therapeutic formula was explored on a trial-error basis, using clinical criteria and *MMPI* total score changes as a basis for either continuing or discontinuing a given treatment strategy. At the present time, laboratory analyses of changes in certain biochemical variables during intensive vitamin therapy are being studied. Exploratory work during the past year appears to indicate that the following variables can either be raised or lowered, depending upon the complex of vitamins and minerals employed: blood glucose, total plasma lipids, total cholesterol, blood pH and the related bicarbonate/carbonic acid ratio, urinary pH, salivary pH, blood pressure, and pulse. Additional variables, particularly the adrenal cortical steroids, are now being studied.

Clinical indication of hypervitaminosis A were not observed, although intakes were high for some subjects. Slight elevations in vitamin A blood level were noted as follows: Subject 7, 130 mcg./100 ml.; Subject 4, 85 mcg./100 ml.; and Subject 1, 105 mcg./100 ml. (normal 20-60 mcg./100 ml.). Subject 7 also had a small elevation in total lipids (865 mg./100 ml., normal 385-840 mg./100 ml.). Prothrombin times were normal for these cases.

It is of interest to note that the average improvement in *MMPI* scores 26.77 points for this study was distributed fairly evenly through the eight scales on the test. This would appear to warrant the interpretation that hypochondriasis, depression, hysteria, psychopathic deviation, paranoia, psychasthenia, schizophrenia, and hypomania, as defined and measured on the Minnesota test, are partly dependent upon the nutritional state of the individual, at least in those instances where improvements are obtained by the therapeutic measures here employed.

F. SUMMARY AND CONCLUSION

Thirty volunteer unhospitalized emotionally disturbed subjects were treated with experimental formulas of large quantities of vitamins and minerals for an average period of 10 months. All subjects were first administered placebos, and were switched to experimental tablets without their knowledge when they failed to improve on placebos. The average length of the placebo trial was 2.46 months. A "total score" on the Minnesota Multiphasic Personality Inventory was computed, and the test was employed for retesting as an objective measure of improvement; in addition, a clinical assessment was made of behavioral changes.

The clinical improvements for the placebo trial were not statistically significant. Of 30 placebo subjects, seven improved, six got worse, and 17 showed no change. Evaluated by changes on the Minnesota Multiphasic Personality Inventory, the average placebo reduction in total test score was 4.40 points. This difference for the placebo trial is not statistically significant.

Clinical evaluations were made at the end of a "first experimental period" as well as at the end of the experiment. The first experimental period averaged 1.76 months. This period of time was determined by matching the lengths of individual placebo trials with as nearly equal periods of experimental trials as possible.

Clinical evaluations at the end of the first experimental period indicated 22 subjects improved, two worse, and six unchanged. This difference is significant well beyond the .01 level of confidence. The average reduction in total test score on the Minnesota Multiphasic Personality Inventory at the end of the first experimental period was 17.13 for 30 subjects. This improvement is significant well beyond the .01 level of confidence.

Clinical evaluations at the completion of the experiment showed 24 subjects improved, one worse, and five unchanged. This difference is significant well beyond the .01 level of confidence. Evaluating the total experimental change in terms of improvement on the Minnesota test, the average reduction in total test score was 26.77 points for 30 subjects (from 219.03 to 192.26), which again is statistically significant well beyond the .01 level of confidence.

Comparing the effect of the total experimental treatment directly with the placebo effect (net change in total score on the Minnesota Multiphasic Personality Inventory on experimentals minus net change in *MMPI* total score on placebos), the mean net score reduction for the experimental treatment was 22.37 points, a difference which is significant beyond the .01 level of confidence.

The experiment here reported is the third in a continuing series of investigations of some of the nutritional aspects of mental illness (7, 8). These studies indicate that some states which are psychologically diagnosed as functional mental illness, and at the same time are not accompanied by clinical evidence of nutritional deficit, apparently involve unsuspected nutritional deficiencies, and may be helped by appropriate nutritional therapy.

REFERENCES

1. BELL, E. C. Nutritional deficiencies and emotional disturbances. *J. of Psychol.*, 1958, **45**, 47-74.
2. BROZEK, J., & ERICKSON, N. K. Item analysis of the psychoneurotic scales of the Minnesota Multiphasic Personality Inventory in experimental semi-starvation. *J. Consult. Psychol.*, 1948, **12**, 403-411.
3. HATHAWAY, S. R., & MEEHL, P. E. An Atlas for the Clinical Use of the *MMPI*. Minneapolis: Univ. Minnesota Press, 1951.
4. KEYS, A., *et al.* The Biology of Human Starvation. Minneapolis: Univ. Minnesota Press, 1950.
5. KJENAAS, N. K., & BROZEK, J. Personality in experimental semi-starvation. *Psychosomat. Med.*, 1952, **14**, 115-128.
6. WELSH, G. S., & DAHLSTROM, W. G., (*Eds.*). Basic Readings on the *MMPI* in Psychology and Medicine. Minneapolis: Univ. Minnesota Press, 1956.
7. WATSON, G. Vitamin deficiencies in mental illness. *J. of Psychol.*, 1957, **43**, 47-63.
8. WATSON, G., & COMREY, A. L. Nutritional replacement for mental illness. *J. of Psychol.*, 1954, **38**, 251-264.

W. C. Kalash Laboratories
1102 East Walnut Street
Pasadena 4, California

SECONDARY REINFORCEMENT WITH AND WITHOUT CONTROL FOR RESPONSE FACILITATION*

Department of Psychology, University of Mississippi

WILLIAM P. WILKES AND WILLIAM F. CROWDER

A. INTRODUCTION

Several writers on Skinner-box studies of secondary reinforcement (e.g., 1, 2) have pointed to a possible artifact, namely, the evoking or facilitating effect of the secondary reinforcing signal. One way to control such evocation is to measure the strength of the lever response in the absence of the signal, on a day subsequent to lever training. This "retention method" has been used in one recent experiment (1).

The purpose of the present study was to compare the retention method with a conventional measure of secondary reinforcement. Two experiments are reported in which differential responding to a pair of levers was studied. One experiment tried to build up a lever preference on the basis of a small number of secondary reinforcements. The other experiment used a somewhat larger number of secondary reinforcements, and included an additional group of Ss receiving primary (food) rewards. Each experiment employed two measures of secondary reinforcement: (a) the number of responses made to each bar *during* lever training; (b) the number of responses made to each bar on the day *following* lever training.

B. EXPERIMENT I

1. Method

a. *Subjects and apparatus.* The Ss were 13 naïve Holtzman albino rats, about 90 days old. The apparatus was eight feeding boxes, four magazine-training boxes, and a two-bar Skinner box. The feeding and magazine boxes are described elsewhere (1). A magazine box contained a buzzer, a food vendor, and a food cup. The Skinner box had a buzzer and two identical bars, 2 in. wide, $3\frac{1}{2}$ in. apart, 4 in. high, and operating on 11 gm.

b. *Procedure.* The Ss were housed one pair to a cage with water always available. The experiment lasted 20 days: 13 days of deprivation; 1 day of magazine training; 1 day of feeding; 3 days of recording operant level;

* Received in the Editorial Office on September 24, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

1 day of lever training via secondary reinforcement; and 1 day of testing for retention of lever training. The *Ss* had been without food for 23 hr. at the start of each session, and were fed at the end of each session.

(1). *Deprivation schedule.* After three days without food, the *Ss* were fed in feeding boxes on each of the next 10 days. The ration consisted of 25 magazine pellets, plus 7 gm. of laboratory food.

(2). *Magazine training.* Each *S* was placed in a magazine box and received 80 pairings of a 1-sec. buzzer with a .05-gm. food pellet. Three pairings were given every 6 min. at unequal intervals.

(3). *Feeding.* Food was provided at the usual eating time in order to re-establish the previous feeding schedule, which had been interrupted by the relatively lengthy magazine-training session.

(4). *Operant level test.* The *Ss* were left in the Skinner box for 30 min. without the buzzer, and responses were recorded.

(5). *Lever training via secondary reinforcement.* Each *S* was placed in the Skinner box and was removed as soon as possible after 10 responses had been reinforced by the 1-sec. buzzer. For six randomly selected *Ss*, the reinforced bar had been preferred on the third operant-level day; for the other seven *Ss*, the previously nonpreferred bar was reinforced.

(6). *Retention test.* The *Ss* were left in the Skinner box for 30 min. and responses to both levers were recorded, as in the operant level test.

2. Results

During lever training, the mean number of responses to the reinforced bar was 11.5; to the other bar, 5.9. This difference was significant at the .02 level, according to the *t* test ($t = 3.02$). On the retention test day, the mean number of responses to the previously reinforced bar was 23.2; to the other bar, 26.2. This difference did not approach significance ($t = -.273$). Thus the differentiation of responding observed in bar training was not retained on the next day, when the buzzer was not presented.

C. EXPERIMENT II

This experiment differed from the first one in several respects: the magazine training occurred in the Skinner box; somewhat more lever training was given; and a primary reinforcement group was included (to check on the "difficulty" of the two-bar task). Also, lever training was terminated after a fixed number of responses had been made on both levers combined. It was thought that the use of a fixed number of reinforced responses in Experiment I might have biased the training-day scores.

1. Method

a. Apparatus and subjects. The apparatus consisted of the feeding boxes and the Skinner box of Experiment I, with a food cup and a vendor added to the Skinner box. The *Ss* were 19 male Holtzman rats, about 90 days old. They had previously been trained in the same Skinner box, in a study of the reinforcing effect of weak light. So presumably the *Ss* had initial biases favoring one or the other of the levers.

b. Procedure. The *Ss* were housed two or three to a cage, with water always available. The experiment lasted 12 days: 9 days of food deprivation, and 1 day each of magazine training, lever training, and testing. All sessions began with *S* under 23 hr. of food deprivation, and ended with *S* having eaten 8 gm. of food.

(1). *Food deprivation schedule.* After three days without food, the *Ss* were fed in feeding boxes for the next six days. The daily ration consisted of 20 magazine pellets, plus 7 gm. of laboratory food.

(2). *Magazine training.* Eighty buzzer-pellet pairings were given in the Skinner box, the first 60 at 20-sec. intervals; the final 20 unequally spaced at three per 6 min.

(3). *Lever training.* Nine of the *Ss* were reinforced by the buzzer alone; 10 were reinforced by the buzzer and a food pellet. About half the *Ss* in each group were reinforced on the left bar; the other half on the right bar. Within these restrictions the assignment of *Ss* to conditions was random. Each *S* was removed as soon as possible after a total of 30 responses had been made on both bars together.

(4). *Retention test.* Each *S* was left in the Skinner box for 30 min. without the buzzer, and the responses to each lever were recorded.

2. Results

a. Secondary reinforcement. During lever training, the mean number of responses made to the reinforced bar was 20.4; to the unreinforced bar, 10.6. The difference was significant at the .001 level, according to the *t* test ($t = 6.66$). During retention, the mean number of responses on the previously reinforced bar was 20.2; on the other bar, 20.1. This difference did not approach significance ($t = .009$).

b. Primary reinforcement. During training, the reinforced bar was pressed a mean of 25.4 times, vs. 4.7 for the other bar. This difference was significant at the .001 level ($t = 11.0$). During retention, the reinforced bar was pressed a mean of 42.3 times, vs. 28.8 for the other bar. The difference was significant at the .05 level ($t = 2.47$).

D. DISCUSSION

In both experiments, the two tests led to opposite conclusions about the reinforcing effect of the buzzer. During lever training, the buzzer appeared to be reinforcing, but the retention test showed the effect to be a transitory one. The retention scores of the primary reinforcement group indicate that the two-bar task is not too difficult if a "strong" reward is given. We would attribute the training-day performance of both secondary reinforcement groups at least partly to the evocation of bar-pressing by the buzzer. Having just pressed the reinforced lever, *S* would be in a position to press that lever at the sound of the buzzer. Had the training-day measures alone been taken, only a few secondary reinforcements would have seemed sufficient to condition a bar-preference. Similar findings are reported by Wyckoff, Sidowski, and Chambliss (2), who controlled facilitation in a different way and used single-bar Skinner boxes.

E. SUMMARY

The purpose of the study was to compare two measures of secondary reinforcement, one of which controlled for facilitation of the response by the reinforcer. Two experiments were performed, both employing a two-bar Skinner box. A buzzer was paired with food and was then presented following responses to one of the bars. In both experiments, more responses were made to the reinforced bar, but this could have resulted from facilitation as well as from learning. To control facilitation, the retention of the preference was measured on the day following lever-training, in the absence of the buzzer. No preference for the previously reinforced bar was shown in the retention tests. It appeared that the buzzer was not a strong enough reinforcer to bring about a learned preference between bars. Another group of *Ss*, which received primary reinforcement (food), did retain the preference for the previously reinforced bar.

REFERENCES

1. CROWDER, W. F., GAY, B. R., FLEMING, W. C., & HURST, R. W. Secondary reinforcement or response facilitation?: IV. The retention method. *J. of Psychol.* 1959, **48**, 311-314.
2. WYCKOFF, L. B., SIDOWSKI, J., & CHAMBLISS, D. J. An experimental study of the relationship between secondary reinforcing and cue effects of a stimulus. *J. Comp. & Physiol. Psychol.* 1958, **51**, 103-109.

Department of Psychology
University of Mississippi
University, Mississippi

STYLISTIC VARIATIONS AMONG PROFESSIONAL RESEARCH SCIENTISTS*^{1, 2}

Institute of Personality Assessment and Research, University of California

HARRISON G. GOUGH AND DONALD G. WOODWORTH

A. PROBLEM

It is a matter of common observation in most fields of endeavor where high-level or professional talents are involved, e.g., in music, writing, artistic creation, athletics, etc., that stylistic differences as well as competence differences exist among practitioners. One would certainly expect to find similar stylistic variations among research scientists, although there seems to be very little research work by psychologists that one can turn to for evidence bearing on this anticipation. In our study at the Institute of Personality Assessment and Research of the research originality and general competence of professional workers in the physical sciences and engineering we therefore decided to attempt an exploration of this facet of the scientist's activity.

B. METHOD

1. *Materials*

The first task was to assemble a list of evaluative and descriptive statements which would encompass the range of approaches to research found among working scientists and which would possess the classificatory subtlety needed to delineate the differences between one man and another. In the spring of 1957 we spent a period of about two months working with consultants from physics³ and several engineering fields writing such material, and seeking to insure the inclusion of items possessing the required degree of relevance and diagnostic clarity. The preliminary list of statements was reviewed by members of the Institute staff and by others, resulting in the

* Received in the Editorial Office on September 28, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This report covers a portion of the work being carried out in a five-year program on the general problem of creative originality. The full program includes individual projects studying creativeness in different fields, such as architecture, writing, the physical sciences, mathematics, etc. The program is being conducted by the University of California Institute of Personality Assessment and Research, Dr. D. W. MacKinnon, Director, under a grant from the Carnegie Corporation of New York.

² A preliminary report on some of the findings in this paper was presented at the April, 1958, meetings of the Western Psychological Association.

³ We wish to thank especially our chief consultant, Dr. Franklin C. Hurlbut, for his help in devising the list of 56 diagnostic items.

end in a set of 56 items suitable for *Q*-sorting or check-list usage, and for use by a scientist to describe his own work or that of another person.

The list of 56 items is presented in Table 1. For our study the items were used in *Q*-sort form. Each one was printed on a separate card with the subjects being asked to sort the deck of cards for personal relevance using the frequencies 5, 12, 22, 12, and 5. The task was found to be interesting by nearly all of the men examined, and typically took only 15 or 20 minutes to complete.

2. Subjects

The sample of subjects was comprised of 45 professional research scientists drawn from three industrial laboratories on the West Coast.⁴ These men were all volunteers for participation in a three-day program of assessment at the Institute in Berkeley, but each man had earlier been designated by the personnel director of his company as being employed in a scientific research capacity. Completion of the "Research Scientist *Q*-sort Deck" (*RSQD*) was one of many assessment tasks each man carried out while he spent his weekend at the Institute.

As might be expected, the sample reveals a preponderance of subjects with advanced educational training. Twenty-eight held Ph.D. degrees, four the M.S. degree, eight had had some graduate work, two held the B.S., whereas three had not completed a college degree. The age distribution was as follows: 25-29—five; 30-34—seventeen; 35-39—fifteen; 40-44—four; 45-49—three; and 50-54—one; the mean age was 35.69, standard deviation 5.46.

On the Terman Concept Mastery Test, Form *T* (Terman, 1956) the mean score was 118.2, standard deviation 29.4. This is fairly high average, as evidenced by these illustrative averages taken from Terman's *Manual: Subjects of the Stanford Gifted Study* $\bar{x} = 136.7$, $SD = 28.5$; applicants for Ford Foundation Fellowships, $\bar{x} = 117.9$, $SD = 35.1$; college graduates, $\bar{x} = 112.0$, $SD = 32.0$; spouses of gifted subjects, $\bar{x} = 95.3$; $SD = 37.0$; and Air Force captains, $\bar{x} = 60.1$, $SD = 31.7$.

3. Procedure

Various analyses of the *Q*-sort protocols could and are being conducted. To serve our present interest in the stylistic variations in approaches to scientific and research activity it was felt that an inverse factor analysis,

⁴ The laboratories included those of the Dalmo-Victor Company, San Carlos, California; the Lockheed Corporation Laboratories, Sunnyvale, California; and Varian Associates, Stanford, California. We are indebted to these research centers for their very great help in the conducting of this project, and to the individual scientists from each laboratory who gave so much of their time and energy to it.

TABLE 1
LIST OF ITEMS IN THE "RESEARCH SCIENTIST Q-SORT DECK"

1. Reacts quickly to research problems; immediately generates a great number of ideas.
2. Is somewhat deficient in his command of basic sources and technical literature in the field.
3. Pursues details and ramifications of research problems with great thoroughness.
4. His research interests lie within a rather narrow range.
5. Has exceptional facility in mathematical analysis.
6. Easily discouraged; needs help and encouragement to do his best work.
7. Is keenly interested in methodological aspects of research.
8. Prefers to work on problems which lend themselves to elegant and exact solutions.
9. Tends to slight the contributions of others; takes undue credit for himself.
10. Likes to play his hunches in research; is guided by his subjective impressions.
11. Is neat and orderly in his habits and manner of work.
12. Dislikes and avoids administrative details connected with research projects.
13. Is good at developing short-cuts and approximation techniques.
14. Frequently makes errors; his work needs to be checked for accuracy.
15. Prefers to work alone; is not a "team" research man.
16. Is erratic in his research output; varies from work of excellent quality to work of marginal or even inferior worth.
17. Indifferent to the practical implications of his own research.
18. Seldom comes up with a really new idea or suggestion.
19. Has exceptionally high standards of research performance for himself as well as for others.
20. Has a special talent for solving instrumentation problems.
21. Is thorough and patient in his approach to research issues; does not get upset if progress is slow.
22. Is a driving, indefatigable research man; cannot stop working on a problem until it is solved.
23. Is relatively uninformed on most subjects other than his research specialty.
24. Good at evaluating research; able to diagnose strong and weak points in a program quickly and accurately.
25. Is fiercely competitive; wants to be the best man in every research task that he undertakes.
26. Seeks out the help and advice of other people when he hits a trouble spot in his own research.
27. Makes a serious effort to read current publications and to "keep up" on the literature in his field.
28. Is creative in anything he tries, whether in science or not.
29. Has an orderly, well-organized approach to research; plans his projects and activities with great care and precision.
30. Gives freely of his own time and ideas to other research endeavors without asking for special credit or recognition.
31. Likes to talk out his research ideas and get other people's reactions.
32. Many of his ideas turn out to be impractical.
33. Is aware of his own professional limitations; does not attempt what he cannot do.
34. Has a quick tempo of thought and speech.
35. Prefers to think in analytical and mathematical ways, rather than in terms of physical and structural models.
36. Is never too busy to "talk shop" with other researchers.
37. Has an exceptionally good memory.
38. Has a knack for improvising quick solutions in research trouble spots.
39. Is a research perfectionist; devotes endless attention to matters of design, apparatus, procedure, etc.
40. Plays his cards "close to the vest"; prefers not to tell anyone about his research plans until his work is finished.

TABLE 1 (*Continued*)

41. Is a creative and inventive researcher.
42. Is intolerant of metaphysical issues.
43. Is a talented "re-write" man; can take other people's ideas and hunches and fashion them into practical research designs and programs.
44. Has strong research biases; is vehement in his disapproval of certain methods and procedures.
45. Primarily an "idea man"; prefers to turn his hunches and hypotheses over to someone else for systematic experimentation and analysis.
46. Somewhat given to bluffing; claims to know more than he does.
47. Stimulating to other people; seems to catalyze others into more original and productive endeavor than they would otherwise achieve.
48. Tends to be sarcastic and disparaging in describing the work of other researchers.
49. Is intellectually gifted.
50. Has a lively sense of intellectual curiosity and inquiringness, a desire to know and to understand.
51. Is flexible and adaptable in his thinking; able to shift and to restructure easily.
52. Takes an esthetic view; is sensitive to matters of form and elegance in research problems.
53. Has a "sense of destiny" with respect to his own research career, an inner conviction of the worth and validity of his own efforts.
54. Enjoys philosophical speculation.
55. Lacks confidence, is afraid to strike out in new directions.
56. Subordinates everything to his research and scientific goals; puts scientific values above all others.

or factor analysis of "persons" would be the method of choice. The specific technique employed was a principal axis factor analysis, programmed for the IBM 701 computing machine by Mr. Jack Neuhaus of the University of California Electronic Computing Center, followed by the varimax rotation for factorial invariance developed and programmed by Dr. Henry Kaiser (1958) of the University of Illinois Bureau of Research and Service.

Because the factor analysis program could accommodate only 36 persons, the calculations were first carried forward on the matrix of the first 36 names in an alphabetical list. From this matrix eight "person-factors" were extracted, which after rotation accounted for 22 per cent, 19 per cent, 14 per cent, 10 per cent, 9 per cent, 9 per cent, 9 per cent, and 7 per cent, respectively, of the analyzed variance.⁵ Approximately 71 per cent of the total variance in the matrix was included in the analyzed variance.

Following this, the three to five scientists (of the 36) having highest load-

⁵ The original correlation matrix for the 45 scientists, along with the tables giving rotated and unrotated factor loadings for the subsample of 36 scientists used in the factor analysis, have been deposited as Document number 6069 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Copies may be secured by citing the Document number and by remitting \$1.25 for photoprints, or \$1.25 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress, Washington 25, D. C.

The distribution of coefficients in the 45x45 matrix varied from $-.20$ to $+.69$, with $Q_1 = +.23$, $Q_2 = +.33$, and $Q_3 = +.41$.

ings on each factor were identified, and then work sheets were set up giving the average zero-order correlation of each of the 45 scientists with the composite of scientists chosen as the "marker" for each type. The purpose of this improvisation was merely to discover whether any of the nine men not included in the factor analysis had prominent enough correlations with any of the eight dimensions to justify being included in the reference or "marker" composite.

The next step was to select from this work sheet of average correlations the cluster of men to be used in deriving the reference *Q*-sort for each of the eight factor dimensions. From three to five men were again chosen to establish each composite. The individual *Q*-sorts for the men chosen were totaled and then re-sorted into *Q*-form. These eight new *Q*-sorts constitute the reference or defining *Q*-sorts for each of the factor types. Table 2 presents these reference *Q*-sorts for each of the eight types.

TABLE 2
REFERENCE *Q*-SORTINGS FOR THE EIGHT STYLISTIC TYPES ON THE RESEARCH SCIENTIST
Q-SORT DECK*

Card	Reference sortings								Card	Reference sortings							
	I	II	III	IV	V	VI	VII	VIII		I	II	III	IV	V	VI	VII	VIII
1.	3	5	4	3	2	4	3	3	29.	3	3	3	4	3	2	2	2
2.	2	3	2	3	4	3	3	4	30.	4	5	3	3	5	4	4	2
3.	4	3	3	3	3	3	3	3	31.	4	4	3	4	4	3	5	4
4.	3	2	2	2	3	1	3	3	32.	2	3	2	3	1	3	3	3
5.	5	2	4	4	2	5	5	2	33.	3	3	4	4	5	3	4	4
6.	2	2	2	2	4	3	2	2	34.	3	4	3	3	2	3	3	4
7.	3	2	3	3	4	3	5	3	35.	4	3	3	3	1	5	5	1
8.	5	3	3	3	3	5	4	3	36.	4	4	3	4	5	4	3	4
9.	1	1	2	2	1	2	1	1	37.	3	3	2	5	1	3	2	3
10.	3	3	3	3	3	4	3	4	38.	3	4	5	3	4	3	2	4
11.	3	3	3	5	3	1	3	2	39.	3	2	3	1	3	1	2	3
12.	4	3	3	1	3	4	3	5	40.	2	2	3	3	3	3	1	2
13.	2	4	4	4	4	4	4	3	41.	5	5	4	3	2	3	3	3
14.	1	2	1	2	2	4	2	3	42.	2	2	4	5	1	2	3	3
15.	3	1	4	3	3	4	3	5	43.	3	3	3	4	5	3	3	3
16.	2	3	1	2	3	3	3	3	44.	3	1	1	1	3	2	1	3
17.	2	2	3	2	3	2	4	3	45.	3	4	3	2	2	3	3	3
18.	1	2	1	3	3	3	2	2	46.	2	3	2	1	2	2	2	2
19.	4	3	4	4	3	4	3	3	47.	3	5	4	3	2	3	4	3
20.	2	4	2	3	4	2	1	4	48.	1	1	1	2	2	1	2	2
21.	3	3	3	4	4	1	3	2	49.	4	3	4	3	3	3	3	4
22.	5	3	3	3	3	4	3	1	50.	5	5	5	3	4	5	4	5
23.	2	3	2	1	2	2	2	3	51.	3	4	4	5	4	4	5	2
24.	4	4	5	4	3	3	4	3	52.	3	3	3	3	4	5	3	4
25.	3	2	4	2	2	3	1	3	53.	4	3	3	2	3	2	3	4
26.	4	4	2	5	4	2	4	4	54.	2	3	5	4	5	4	3	5
27.	4	4	3	4	3	2	4	4	55.	1	1	2	3	3	2	2	1
28.	3	4	5	2	3	3	4	1	56.	3	2	2	2	2	3	2	2

* The higher the numerical value, the more salient or characteristic is the phrase indicated by the number of the card.

The intercorrelations among the eight composites are presented in Table 3. The range of coefficients is from $-.32$ to $+.51$, with a median r of $+.12$. These values seem low enough to warrant the inference that eight separate stylistic types are being defined by the scoring criteria (reference sorts).

TABLE 3
INTERCORRELATIONS AMONG THE SCORING KEYS FOR THE EIGHT STYLISTIC TYPES ON THE
RESEARCH SCIENTIST *Q*-SORT DECK

	I	II	III	IV	V	VI	VII	VIII
I.	—							
II.	+.21	—						
III.	+.47	+.32	—					
IV.	+.13	+.25	+.25	—				
V.	-.30	+.13	-.22	+.27	—			
VI.	+.30	-.23	+.03	-.29	-.32	—		
VII.	+.51	+.12	+.27	+.38	.00	+.19	—	
VIII.	-.23	+.12	-.12	-.13	+.26	-.16	-.25	—

In order to obtain a profile of "stylistic type" scores for each man, the individual *Q*-sorts of each of the 45 scientists were correlated with each of the eight reference sorts. (Mean coefficients for each of the eight reference dimensions were: .46, .45, .45, .37, .28, .32, .41, and .32). The correlation between the scientist's *Q*-sort and the reference sort was treated as his score on that factor. The 45 values for each factor dimension were then correlated with scores on other devices and with the ratings and evaluations made of each subject by the staff interviewers and observers.

C. DETERMINATION OF THE PSYCHOLOGICAL MEANING OF THE FACTORS

The next phase of the analysis was an attempt to determine the psychological meaning implicit in each of the eight stylistic types. To prepare such formulations a detailed study was made of the order of placement of all 56 cards in the specific composite, and also of the correlations between the "scores" described in the preceding paragraph and the other assessment data. The evidence here included various aptitude and achievement tests, situational procedures, psychodiagnostic life history interviews, experimental tasks, reports and comments made by staff observers, etc.

An example of one of the latter sources of data is a 50-card "interpersonal *Q*-sort deck" prepared by the senior author for use in describing the everyday, interpersonal behavior of a subject. This *Q*-sort deck was sorted in description of each of the 45 subjects by three persons: the psychologist who had seen him in the life history interview and two staff psychologists who had observed him during the various procedures and inter-

actions of the three-day assessment program. Each of the eight sets of stylistic type scores was correlated with each of these 50 descriptive statements. Representative items revealing highest correlations with each stylistic type are listed below:

Highest Positive and Negative Correlations

Factor Type I. Positive correlations: Tolerant, permissive, and benevolent; considerate and charitable (+.36). Has highly developed inner sense of ethics and morality; deeply humanitarian and altruistic (+.35). A conscientious and serious-minded person (+.25). *Negative correlations:* Wedded to routine; made anxious by change and uncertainty (—.28). Gets along well with others; able to "fit in" easily in most situations (—.43).

Factor Type II. Positive correlations: Ambitious, likely to succeed in most things undertaken (+.43). Well-organized, capable, patient, and industrious; values achievement (+.40). Efficient; able to mobilize personal resources quickly and effectively (+.32). *Negative correlations:* Lazy, indifferent about duties and obligations; generally undependable and immature (—.50). Worried and preoccupied; tense, nervous, and generally upset (—.53).

Factor Type III. Positive correlations: Is an effective leader; able to elicit the response and cooperation of others (+.31). Is forceful and self-assured in manner (+.26). Poised and self-confident; not troubled by pressure or criticism (+.26). *Negative correlations:* Shrewd and self-centered; inclined to be selfish and opportunistic (—.36). Awkward and ill-at-ease socially; shy and inhibited with others (—.28).

Factor Type IV. Positive correlations: Is a conscientious and serious-minded person (+.33). Easily embarrassed; feels inferior and inadequate (+.26). *Negative correlations:* Lazy, indifferent about duties and obligations; generally undependable and immature (—.33). Active and robust in manner; hard-headed and forthright in judgment (—.29). Coarse and vulgar; apt to behave in a crude and impolite fashion (—.28).

Factor Type V. Positive correlations: Honest and direct in behavior; mature and realistic in outlook (+.39). Gets along well with others; able to "fit in" easily in most situations (+.38). Observant and perceptive; quick to respond to the subtleties and nuances of others' behavior (+.37). *Negative correlations:* Given to moods; and resentful of others; lack-lustrant (—.40). Headstrong, rebellious, and resentful of others; lacking in self-discipline; apt to behave in a rash or destructive manner (—.38).

Factor Type VI. Positive correlations: Lazy, indifferent about duties and obligations; generally undependable and immature (+.50). Undependable, poorly motivated; has difficulty in working toward prescribed goals (+.48). Clever and imaginative; a spontaneous and entertaining person (+.35). *Negative correlations:* Well-organized, capable, patient,

and industrious; values achievement ($-.59$). Wedded to routine; made anxious by change and uncertainty ($-.53$).

Factor Type VII. Positive correlations: Tolerant, permissive, and benevolent; considerate and charitable ($+.32$). Has a talent for creative and original thinking ($+.29$). Given to moods; often difficult and recalcitrant ($+.28$). *Negative correlations:* Overly concerned with success; too dominated by own ambition and desire to win approval ($-.38$). Warm and unpretentious in behavior; a comfortable and uncomplicated person ($-.31$).

Factor Type VIII. Positive correlations: Active and robust in manner; hard-headed and forthright in judgment ($+.41$). Coarse and vulgar; inclined to behave in a crude and impolite fashion ($+.31$). Overly concerned with success; too dominated by own ambition and desire to win approval ($+.28$). *Negative correlations:* Worried and preoccupied; tense, nervous, and generally upset ($-.29$). Self-pitying and martyr-like; tries to make others feel guilty ($-.29$).

A second example of the kind of data utilized in reaching a formulation of the psychological meanings of each of the eight stylistic factors is taken from the personality inventory scores on the California Psychological Inventory (Gough, 1957). For brevity's sake, we are citing the findings only for the second factor.

TABLE 4
COMPARISON ON THE CALIFORNIA PSYCHOLOGICAL INVENTORY OF THE 13 HIGHEST-SCORING VS. THE 13 LOWEST-SCORING SCIENTISTS ON THE SECOND STYLISTIC FACTOR

CPI Scale	Highs		Lows		diff.	t	P
	M	SD	M	SD			
Dominance	33.69	3.40	30.31	4.66	3.38	2.11	<.05
Capacity for status	24.31	2.63	23.31	2.29	1.00	1.03	*
Sociability	27.00	4.69	27.85	3.80	-0.85	0.51	*
Social presence	39.92	5.77	41.85	4.00	-1.93	0.99	*
Self-acceptance	23.46	3.93	24.23	3.09	-0.77	0.55	*
Sense of well-being	40.77	2.39	38.85	2.67	1.92	1.94	*
Responsibility	34.62	3.55	32.62	3.07	2.00	1.54	*
Socialization	39.38	3.48	32.85	3.78	6.53	4.60	<.01
Self-control	34.23	4.80	29.62	5.25	4.61	2.34	<.05
Tolerance	27.54	3.57	25.08	3.75	2.46	1.71	*
Good impression	19.92	4.50	16.46	5.36	3.46	1.78	*
Communality	25.92	1.12	25.46	1.39	0.46	0.92	*
Achievement via conformance	31.46	3.36	29.62	3.04	1.84	1.46	*
Achievement via independence	24.62	2.33	25.08	3.38	-0.46	0.40	*
Intellectual efficiency	43.46	3.07	43.62	2.99	-0.16	1.34	*
Psychological-mindedness	15.46	3.07	16.62	2.18	-1.16	1.12	*
Flexibility	11.54	4.12	13.92	4.13	-2.38	1.47	*
Femininity	16.54	2.44	17.69	1.97	-1.15	1.32	*

* $P > .05$.

Table 4 contains *CPI* data contrasting the scores of the 13 scientists having highest scores on Factor II with the 13 having lowest scores. Three of the differences seem to merit attention: the scientists high on the factor appear to be more dominant and forceful, less given to aggressive and rebellious behavior, and less governed by errant or erratic impulse.

These personality inventory findings are in accord with the previous interpersonal *Q*-sort data which showed Type II scientists to be ambitious, well-organized, patient, and generally free from problems of immaturity and nervous tension. Furthermore, in their own self-views of research activity (see Tables 1 and 2) they report that they react quickly to research problems, immediately generating a great number of ideas, that they give freely of their own time, that they consider themselves to be creative and inventive, and that they have a lively sense of intellectual curiosity. Utilizing these and other data in the assessment files we therefore decided to designate the Type II scientist as the "initiator."

D. PSYCHOLOGICAL DEFINITIONS OF THE RESEARCH TYPES

Our search for names and formulations for each of the eight types followed the lines suggested in the above material. It was systematic in that all of the available data (including clinical and life history data) were considered, and painstaking in that many hours of analysis were spent on the task. However, it has to be admitted that there is a great and imponderable element of subjectivity in the endeavor and that the formulations represent inductions arrived at by reflective and psychological processes rather than by means of formal statistical analysis. Our goal has been to delineate a brief but integrated psychological definition of the scientist typified by each of the eight stylistic factors; it is hoped that the formulations below represent at least partially valid steps in this direction.

Type I: The Zealot. This man is dedicated to research activity; he sees himself as a driving, indefatigable researcher, with exceptional mathematical skills and a lively sense of curiosity. He is seen by others as tolerant, serious-minded, and conscientious, but as not getting along easily with others and as not being able to "fit in" readily with others.

Type II: The Initiator. This man reacts quickly to research problems, and begins at once to generate ideas; he is stimulating to others and gives freely of his own time; he sees himself as being relatively free of doctrinaire bias—methodological or substantive—and as being a good "team" man. Observers describe him as ambitious, well-organized, industrious, a good leader, and efficient. They also characterize him as being relatively free of manifest anxiety, worry, and nervousness.

Type III: The Diagnostician. This man sees himself as a good evaluator, able to diagnose strong and weak points in a program quickly and

accurately, and as having a knack for improvising quick solutions in research trouble spots. He does not have strong methodological preferences and biases, and tends not to be harsh or disparaging towards others' mistakes and errors. Observers see him as forceful and self-assured in manner, and as unselfish and free from self-seeking and narcissistic striving.

Type IV: The Scholar. This man is blessed with an exceptional memory, and with an eye for detail and order. However, he is not a research perfectionist nor an endless seeker for ultimates. He does not hesitate to ask help when blocked in his work, and feels that he can adapt his own thinking to that of others. He is well-informed in his field, and is not given to bluffing. Observers describe him as conscientious and thorough, and as very dependable, but as lacking confidence and decisiveness of judgment.

Type V: The Artificer. This man gives freely of his own time, and enjoys talking shop with other researchers. He is aware of his own limitations and does not attempt what he cannot do. He sees himself as having a special facility for taking inchoate or poorly formed ideas of others and fashioning them into workable and significant programs. Observers see him as honest and direct, getting along well with others, and as usually observant and perceptive and responsive to nuances and subtleties in others' behavior.

Type VI: The Esthetician. This man favors analytical over other modes of thinking, and prefers research problems which lend themselves to elegant and formal solutions. His interests are far-ranging, and he tends to become impatient if progress is slow or if emphasis must be put upon orderliness and systematic detail. His own view of experience is primarily an esthetic one. Observers see him as clever and spontaneous, but as undependable and immature, somewhat lacking in patience and industry and indifferent about duties and obligations.

Type VII: The Methodologist. This man is vitally interested in methodological issues, and in problems of mathematical analysis and conceptualization. He is open about his own research plans and enjoys talking about them with others. He has little competitive spirit and tends to take a tolerant view of research differences between himself and others. Observers characterize him as a considerate, charitable person, free from undue ambition; at the same time they report a certain moodiness and an occasional tendency toward complicated and difficult behavior.

Type VIII: The Independent. This man eschews "team" efforts, and dislikes and avoids administrative details connected with research work. He is not a driving, energetic research man, although he does have a lively sense of intellectual curiosity. He prefers to think in reference to physical and structural models rather than in analytical and mathematical ways. Observers describe him as active and robust in manner and hard-headed and forthright in judgment. He appears relatively free from worry and self-doubt, but inclined to behave impolitely or abruptly.

In Table 3 we presented the intercorrelations among the scoring keys for the eight stylistic types. One might well ask what would be the intercorrelations among the scores obtained by the sample of 45 scientists, and also whether this pattern of intercorrelations would be upheld in another sample. Table 5 contains data relevant to an answer to this question. The top line of correlations presents the results obtained with the initial sample of 45 professional scientists. The *RSQD* (Research Scientist *Q*-sort Deck) was also given to a sample of 40 honor students in the School of Engineering at the University of California, Berkeley.⁶ The *Q*-sorts of each student were correlated with the stylistic-type scoring keys, yielding eight coefficients or "scores" for each student. (Mean coefficients for the eight factor dimensions were: .35, .37, .30, .31, .24, .26, .29, and .26.) The 40 scores (coefficients) for each dimension were in turn intercorrelated, with the results shown in Table 5.

TABLE 5
INTERCORRELATIONS AMONG SCORES ON THE EIGHT STYLISTIC TYPES OF THE RESEARCH SCIENTIST *Q*-SORT DECK*

		I	II	III	IV	V	VI	VII	VIII
I.	The Zealot	—							
II.	The Initiator	.47	—						
III.	The Diagnostician	.64	.68	—					
IV.	The Scholar	.47	.32		—				
V.	The Artificer	.49	.44	.48		—			
VI.	The Esthetician	.13	.25	.25	—		—		
VII.	The Methodologist	— .25	— .08	— .24	— .30	—			
VIII.	The Independent	— .30	— .13	— .22	— .27	— .13	—		
		.19	— .02	.24	— .16	— .32	—		
		.30	— .23	.04	— .29	.11	.45	—	
		.59	.24	.44	.54	.11	.19	—	
		.51	.12	.27	.38	.00	.09	.13	—
		— .02	.36	.17	.17	.32	.06	— .25	—
		— .23	.12	— .12	— .12	.26	— .16		

* Correlations in the first row are from the sample of 45 professional research scientists, in the second row from the sample of 40 honor students in engineering.

For the professional scientists, the range of coefficients is from $-.30$ to $+.68$, median $+.12$. For the honor students in engineering, the range is from $-.30$ to $+.47$, median $+.12$. These values are quite similar. The correlation between the two matrices is approximately $+.72$. The general pattern of relationship among the eight types would therefore appear to be relatively stable.

⁶ We wish to thank the 40 students who participated in the project, and also Professor W. Goldsmith and Dean M. O'Brien of the School of Engineering for their part in making this testing possible.

E. SUMMARY

A set of 56 diagnostic statements pertaining to scientific activity, values, and modes of research activity was created and then given to a sample of 45 professional research scientists for self-descriptive *Q*-sorting. A factor analysis of the intercorrelations among scientists permitted the identification of eight factor types. The differences among the types seemed to be principally stylistic ones, and to warrant the conclusion that stylistic variations in scientific research methodologies do exist and can be specified.

REFERENCES

1. GOUGH, H. G. Manual for the California Psychological Inventory. Palo Alto, California: Consulting Psychologists Press, 1957.
2. KAISER, H. F. The varimax criterion for analytical rotation in factor analysis. *Psychometrika*, 1958, **23**, 187-200.
3. TERMAN, L. M., *et al.* Concept Mastery Test. New York: Psychological Corporation, 1956.

Institute of Personality Assessment and Research
University of California
2240 Piedmont Avenue
Berkeley 4, California

STUDIES IN PHYSIOGNOMIC PERCEPTION: VI. DIFFERENTIAL EFFECTS OF DIRECTIONAL DYNAMICS OF PICTURED OBJECTS ON REAL AND APPARENT MOTION IN ARTISTS AND CHEMISTS*¹

Department of Psychology, Clark University

PETER E. COMALLI, JR.

A. INTRODUCTION AND PROBLEM

Perceptual objects are characterized not only by their structural aspects, i.e., geometry of form, extensity, intensity, etc., but also by their "expressive" or dynamic qualities. For example, colors are not only experienced in terms of their hue, brightness, and saturation, but also in terms of being strong or weak, gay or sad. Lines may be seen as cruel or happy; forms as static or active. In contrast to most psychologists who have restricted their considerations to the structural qualities, the distinction between expressive and structural qualities has been treated extensively by Heinz Werner, who has referred to the two modes of cognition as the geometric-technical on the one hand, and as the expressive or physiognomic mode of cognition, on the other hand (8, 9).

At Clark, a series of experiments conducted within the framework of sensory-tonic field theory have been concerned with one aspect of physiognomic perception, namely, directional dynamics. Briefly, this term refers to the fact that some objects seem to express a vectorial quality describable in terms of direction and force. For example, a picture of a bird in flight conveys motion in a particular direction. These experiments have demonstrated that directional dynamics affect the localization of stationary objects as measured by displacements of the apparent median plane (straight-ahead) and the apparent horizon (eye-line) (2, 7). The stimulus materials used were pictured objects and words conveying directional dynamics. Similarly, the effect of directional dynamics was also studied in regards to autokinetic

* Received in the Editorial Office on October 2, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This investigation was supported by research grant M 348 from the National Institute of Mental Health of the National Institutes of Health, Public Health Service.

This study is based on a dissertation submitted to the Department of Psychology, Clark University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author wishes to acknowledge his appreciation to his Doctoral Committee, Heinz Werner, Seymour Wapner, and Gordon Gwinn, for their assistance and guidance.

motion: autokinetic motion occurs consonant with the direction of the dynamics in the stimulus figure (1). More recently, auditory stimulation with a vectorial quality induced by ascending and descending gliding tones was also shown to significantly affect the direction of autokinetic motion (5).

Such physiognomic qualities can be studied and objectively demonstrated by two general methods: through variation of conditions inherent in the stimulus situation (as utilized in the previous studies) or through variation of more permanent conditions inherent in the habitual way in which individuals perceive the world of objects. Thus, if individuals characteristically differ in terms of their orientation to the world of objects, one person more susceptible to the geometric-technical mode of cognition, and the other to the expressive mode of cognition, then one should expect these differences to express themselves in the efficacy of directional dynamics as measured behaviorally. One such study by Wapner and Werner utilizing age groups from 6 to 19 tested the assumption that children are more physiognomically oriented than adults (6); they found that in judging the speed of moving figures the youngest age group perceived dynamic figures as moving faster than static figures and that this effect decreases considerably with increase in age.

At the adult level individuals who are normally assumed to differ characteristically in their apprehension of the world are individuals classified as artists and scientists. Many investigators of artistic creativity and aesthetic expression have remarked on the difference between an aesthetic attitude and a scientific attitude (3, 4, 9). Artists are assumed to be more oriented in terms of the expressive qualities of objects whereas scientists are assumed to be more oriented towards the realistic, matter-of-fact, geometric-technical qualities of objects. This led to the hypothesis then that artists should be more greatly affected by physiognomic qualities than scientists.

Two experiments were designed to test this hypothesis: the first experiment deals with the differential effects of directional dynamics on the apparent speed of real motion; the second with its effect on autokinetic motion.

B. EXPERIMENT I: DIFFERENTIAL EFFECTS OF DIRECTIONAL DYNAMICS OF PICTURED OBJECTS ON APPARENT SPEED OF REAL MOTION IN ARTISTS VERSUS SCIENTISTS

In this experiment the efficacy of directional dynamics was measured by its effect on the apparent speed of moving pictorial objects. It is expected that directional dynamics will have a greater effect on the perception of apparent speed for artists than for the scientists.

1. Procedure

The figures and methodology used were similar to that of a previous study employing children (6). The experimental situation was arranged so that *S* saw two series of moving figures which could be viewed through two windows, 2 in. high and 6 in. wide, cut in a large cardboard screen (5 ft. high and 6 ft. wide). The windows were at eye level, illuminated evenly from behind the screen, and 30 in. apart horizontally. Behind each window a series of identical pictures were drawn 1 in. apart on white paper which served as moving belts. The paper belts were driven by two variable speed motors² which could be independently controlled by the *S*. The *S* was seated 8 ft. from the screen.

Three pairs of figures were used (see Figure 1): (a) a triangle with apex pointing in the direction of motion vs. triangle with apex pointing opposite the direction of motion; (b) a running horse vs. a grazing horse; (c) a running boy vs. a standing boy. For each pair of figures, the static figure³ once served as the standard, and the dynamic figure once served as the standard. The dynamic figure was on the left in all comparisons for both horse and boy figures, whereas the static triangle was on the left. The direction that the pictures moved was always from left to right. The speed of the standard was 8.0 on the variac scale, equivalent to 13.3 cm./sec. *S* by manipulating the speed control adjusted the speed of the variable figure so that for him it appeared to be moving at the same velocity as the standard figure. Four trials were carried out for each pair of figures. On two trials the initial setting of the variable was slow (setting of 6.0 on the variac; equivalent to 9.6 cm./sec), on two trials the initial setting of the variable was fast (setting of 10.0 on variac; equivalent to 17.0 cm./sec.).

The sequence of figures presented were in the order triangle, horse, boy. Half the *S*'s received this sequence first with dynamic figure as standard followed by the static figure as the standard. Half the *S*'s had the reverse.

In sum, there was a total of six test conditions, i.e., judgments of each of three pairs of figures—triangle, horse, and boy—under two different conditions: (a) where the speed of the dynamic figure was adjusted so that it appeared equal to the static; and (b) where the speed of the static figure

² The speed of the motors (1/15 H.P. D-C Shunt Motors) were varied by a Type 1701-AK Variac Speed Control which provides adjustable constant-speed operation of *d-c* motors from *a-c* lines. They were purchased from General Radio Company, Cambridge, Massachusetts.

³ For purposes of convenience only, the triangle pointing opposite the direction of motion is referred to hereafter as a static figure, though it is assumed that it has dynamic properties with its dynamics in the direction of the pointedness of the triangle (opposite the direction of motion in this case).

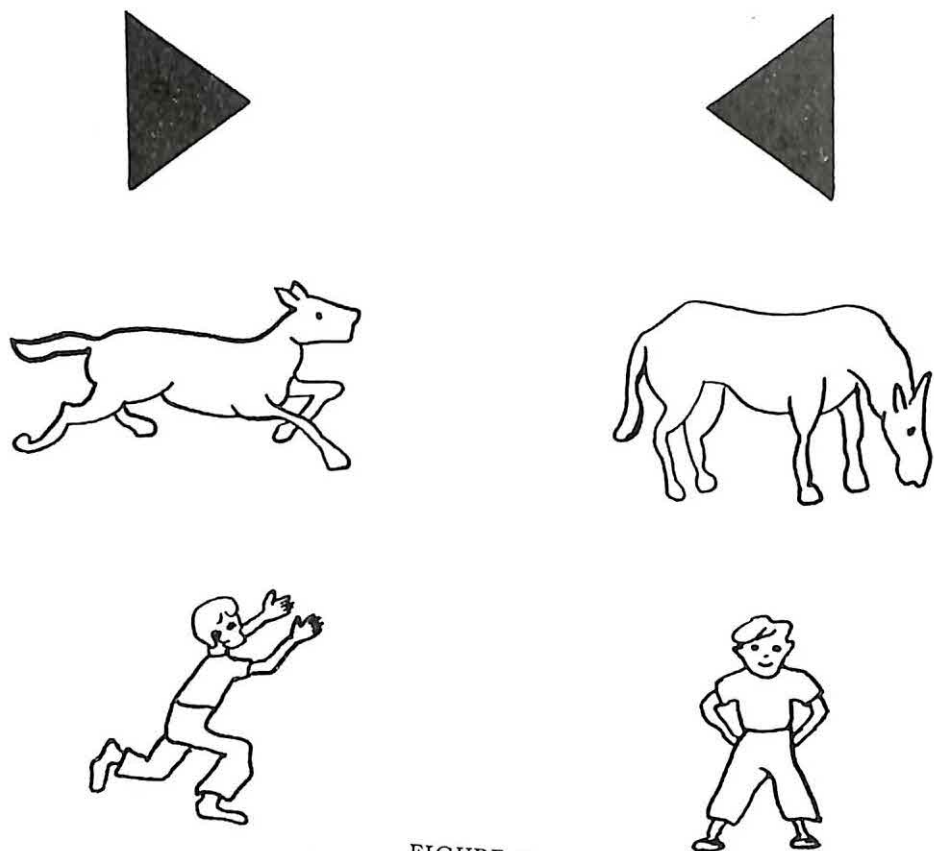


FIGURE 1
EXPERIMENT I: TEST PAIRS

was adjusted so that it appeared equal to the dynamic. The four trials (two with slow and two with fast initial setting of variable) within each of these conditions was averaged and served as the unit of analysis in the experiment.

The data were analyzed by a repeated measurement for independent group analysis of variance including the following classifications: groups (artists vs. chemists), figure adjusted (static, dynamic), and pairs of figures (triangle, horse, boy).

Two groups of persons were selected to represent individuals with habitual ways of viewing the world, namely, chemists and artists. The chemists were composed of graduate students in chemistry at Clark University; the artists were seniors or graduates from the Worcester Museum School of Art. The median age for both groups was 25 years of age. In all, 32 Ss were tested, 16 chemists and 16 artists.

The measure employed was the physical speed of the adjusted figure. Since

the standard was always set at a variac reading of 8.0 (13.3 cm./sec.) the adjustment of the dynamic figure set at slower speeds (readings of less than 8.0) means that it is seen as perceptually faster; the adjustment of the static figure set at higher speeds (readings greater than 8.0) means that it is seen as perceptually slower.

2. Results

The results are presented in Table 1. The table may first be examined with respect to the general effect of dynamics independent of the two groups studied. In keeping with the hypothesis, when the dynamic figures are adjusted, they are set at physically lower speeds (7.83) than the static figures (8.13). This difference is significant beyond the .01 level of confidence.

TABLE 1
EXPERIMENT I: DIFFERENTIAL EFFECT OF DIRECTIONAL DYNAMICS ON THE APPARENT SPEED OF REAL MOTION IN ARTISTS AND CHEMISTS

Source of variation	df	Mean square	F	P
Between individuals (I)	31	.37	—	—
Groups (G)	1	.02	<1.00 ^a	>.05
I's within G's	30	.38	—	—
Within I's	160	.31	—	—
Figure adjusted (A)	1	4.35	27.19 ^b	<.01
Figure pair (F)	2	.01	<1.00 ^c	>.05
A x F	2	3.74	24.93 ^d	<.01
G x A	2	2.50	15.62 ^b	<.01
G x F	1	.02	<1.00 ^c	>.05
G x A x F	2	.48	3.20 ^d	<.05
Pooled I's x AF comb.	150	.16	—	—
Pooled I's x A	30	.34	—	—
Pooled I's x F	60	.15	—	—
Pooled I's x (A x F)	60	.32	—	—
Total	191			

^a Tested against "I's within G's"

^b Tested against "Pooled I's x A"

^c Tested against "Pooled I's x F"

^d Tested against "Pooled I's x (A x F)"

	Means (variac readings)*		Artists	
	Chemists		Dynamic adjusted	Static adjusted
Figure pair	Dynamic adjusted	Static adjusted	8.03	7.88
Triangle	8.11	7.89	7.44	8.55
Horse	7.76	8.23	7.68	8.31
Boy	7.98	7.94	7.72	8.25
Overall means	7.95	8.02		

* The standard was set at a variac reading of 8.0 (13.3 cm./sec.). The adjustment of the dynamic figure set at slower speeds (readings less than 8.0) means that it is seen as perceptually faster than the static figure; the adjustment of the static figure at higher speeds (readings greater than 8.0) means that it is seen as perceptually slower than the dynamic figure.

With respect to the main problem, differences between artists and chemists, the pertinent result pertains to the interaction between groups (artists vs. chemists) and adjustment (dynamic vs. static). The difference in settings for dynamic vs. static figures is expected to be greater for artists than for chemists. The differences in adjustment of the dynamic vs. the static figure for artists, .53, is significantly greater than the difference in adjustment for chemists, .07. This difference is significant beyond the .01 level of confidence. The dynamic figures are seen as moving perceptually faster for the artists than for the chemists.

In addition it may be noted that there is some evidence for differential effect in potency of the different figures used. There is a greater effect for the horse than for the other two figures.

C. EXPERIMENT II: DIFFERENTIAL EFFECTS OF DIRECTIONAL DYNAMICS OF PICTURED OBJECTS ON AUTOKINETIC MOTION IN ARTISTS VERSUS CHEMISTS

In this experiment the efficacy of directional dynamics was measured by its effect on autokinetic motion. It is expected that the efficacy of the vectorial quality of dynamic objects should be greater for a group of artists than for a group of chemists.

1. Procedure

The figures and methodology used were similar to that of a previous study (1). Each of three figures having left-right directional dynamics were presented in a dark room. The figures used were silhouettes of a running horse, a running boy, and an arrow (see Figure 2). Each stimulus figure was presented in two orientations: once facing to the right and once facing to the left. The stimulus figures, approximately 12 sq. inches in area, were attached to the milkglass face of a lightbox and were dimly illuminated by use of a seven-watt bulb operating at 30 volts *AC*. This arrangement gave *S* the impression of a luminous silhouette suspended in space.

Each *S* was led blindfolded into the dark room, seated in a chair with the head fixed into position by a headrest. The stimulus figure was located at *S*'s eye level, straight ahead at a distance of 12 ft. from *S*. Each figure was presented for 30 seconds in addition to the time taken until movement was first reported. During the 30-second time interval following the first report of movement the *S* had to continuously report the direction of movement. The following measures were used: (*a*) direction of the first report of motion; (*b*) duration of motion in a particular direction during the 30-second interval, and (*c*) latency. The duration measure was recorded by use

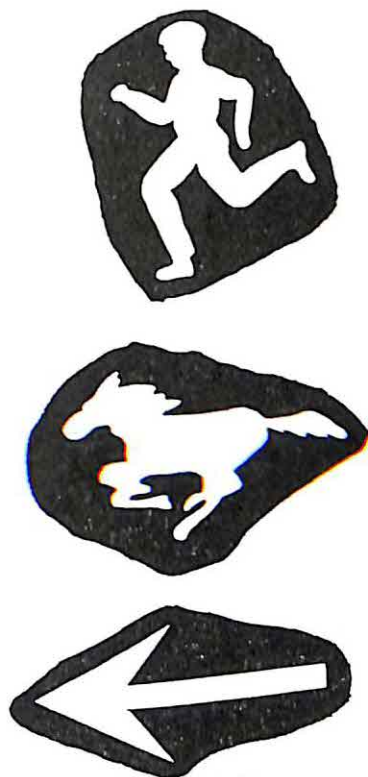


FIGURE 2
EXPERIMENT II: TEST FIGURES

of four clocks. One clock was used for recording movement to the left; another for movement to the right; a third for movement up; and a fourth for movement down. Combined movements such as up-right were recorded by the use of two clocks.

The *S*'s comprising the two groups, chemists and artists, used in the first experiment also served as subjects in this experiment. In all, 32 *S*'s were tested, 16 chemists and 16 artists. A systematic randomized presentation of trials was employed. The data were analyzed as a repeated measurement design with independent groups.

In addition, following the experiment proper, further tests were made with both groups on a non-dynamic stimulus figure, namely a $3\frac{1}{2}$ in. square.

2. Results

The major measure used for the analysis of the data is the difference in duration of autokinetic motion to the right and the duration of autokinetic motion to the left; arbitrarily, a plus score was assigned to indicate pre-

dominance of autokinetic motion to the right, a minus score indicates a predominance of autokinetic motion to the left. Because of inhomogeneity of variance a square root transformation of the difference scores was performed. The analysis and means of these difference scores is shown in

TABLE 2
EXPERIMENT II: DIFFERENTIAL EFFECTS OF DIRECTIONAL DYNAMICS ON AUTOKINETIC MOTION IN ARTISTS AND CHEMISTS

Source of variation	Difference score (R-L)		F	P
	df	Mean square		
Between individuals (I)	31	20.00	—	—
Between groups (G)	1	2.59	<1.00	>.05
I's within G's	30	20.58	—	—
Within I's	160	15.83	—	—
Figures (F)	2	5.30	<1.00	>.05
Dynamics (D)	1	569.59	28.94	<.01
F x D	2	2.92	<1.00	>.05
G x F	2	0.83	<1.00	>.05
G x D	1	25.45	1.29	>.05
G x F x D	2	4.36	<1.00	>.05
Pooled I's x (F x D)	150	10.62	—	—
I's x F	60	8.20	—	—
I's x D	30	19.68	—	—
I's x F x D	60	8.50	—	—
Total	191	14.84	—	—

Mean difference scores (secs.)*

	Chemists		Artists	
	Dynamics left	Dynamics right	Dynamics left	Dynamics right
Horse	—10.3	+8.4	—13.1	+ 7.4
Boy	— 6.6	+7.7	— 8.6	+ 9.8
Arrow	— 5.6	+3.6	— 8.9	+10.4
Overall mean	— 7.5	+6.5	—10.2	+ 9.2

* A plus score indicates predominance of autokinetic motion to the right; a minus score indicates predominance of autokinetic motion to the left.

Table 2. The *F*-test for the general effect of directional dynamics independent of the two groups studied is significant beyond the .01 level of confidence. The overall means (left dynamics = —8.9 seconds; right dynamics = +7.9 seconds) substantiate previous findings that autokinetic motion occurs predominantly in the direction of the dynamics.

With respect to the main hypothesis studied in this experiment, namely, differential effects of dynamics for artists versus chemists, examination of Table 2 shows that there are trends in the expected direction. With dynamics left, the mean difference score for artists is —10.2 seconds; while for the chemists it is —7.5 seconds; with dynamics right, the mean difference score is 9.2 seconds for the artists and 6.5 seconds for the chemists. Al-

though these differences are not statistically significant, this relationship holds in five out of the six test conditions. This same tendency is also indicated by direction of the first motion response in relation to the dynamics in the figure. Inspection of Table 3 shows that for each condition, there is a tendency for a greater number of artists to perceive the first direction of

TABLE 3
EXPERIMENT II: DIFFERENTIAL EFFECTS OF DIRECTIONAL DYNAMICS ON AUTOKINETIC
MOTION IN ARTISTS AND CHEMISTS
(Number of subjects initially seeing left, right, or up-down motion)

Figure	Direction of dynamics	Left		Response category		Up-down	
		Chemists	Artists	Chemists	Artists	Chemists	Artists
Horse	Left	8	12	3	2	5	2
Boy	Left	7	8	3	3	6	5
Arrow	Left	10	10	3	2	3	4
Horse	Right	2	4	6	10	8	2
Boy	Right	1	3	8	9	7	4
Arrow	Right	4	2	8	10	4	4

motion in accordance with the dynamics than for the chemists. These trends lend partial support to the hypothesis that the artists are more strongly affected by directional dynamics than are the chemists. It may simply be noted here, that no differences were found with respect to motion up vs. down in regard to the effect of left-right directional dynamics.

It will be recalled that following the experiment proper, some further tests were made utilizing a figure that has no directional dynamics, namely, a square. Differences between artists and chemists were evaluated here by employing the latency measure. It may be assumed that dynamic figures will initiate motion sooner than a static figure without directional dynamics. Furthermore, if artists are more sensitive to directional dynamics, latency should be shorter for artists than for chemists. An analysis was carried out by computing a mean latency score for the six dynamic figures and comparing this mean latency score with the latency for the non-dynamic stimulus. Because of inhomogeneity of variance a square root transformation of the two mean latency scores was performed and utilized in a repeated measurement, independent group analysis of variance. The *F*-tests and means are given in Table 4. As expected, the overall mean latency for the dynamic figures is significantly shorter, 13.1 secs. than for the static figure, 23.2 secs. Furthermore, the significant interaction of stimulus figure (dynamic, static) and groups (artists, chemists) indicate differential susceptibility to dynamics in the two groups: the mean difference between dynamic and static figures is 14.5 secs. for the artists as compared to 5.3 secs. for the chemists.

TABLE 4

EXPERIMENT II: DIFFERENTIAL EFFECTS OF DIRECTIONAL DYNAMICS ON AUTOKINETIC MOTION IN ARTISTS AND CHEMISTS

Source of Variation	Latency time		F	P
	df	Mean square		
Between individuals	31	4.90	—	—
Between groups (G)	1	.18	<1.00	>.05
I's within G's	30	5.06		
Within Individuals	32	1.79	3.83	<.01
Figures (F)	1	12.07	—	—
G x F	1	5.64	9.14	<.01
Residual	30	1.32	4.27	<.05
Total	63	3.32	—	—

Mean latency time (secs.)			
	Dynamic figure	Square figure	Overall mean
Chemists	15.3		
Artists	11.1	20.8	18.0
Overall mean	13.1	25.6	18.4
		23.2	

D. CONCLUSIONS

The major finding of this study that pictured objects having directional dynamics more potently affects both autokinetic and real motion with a group of artists than with a group of scientists (chemists) supports the proposition that sensitivity to expressive qualities is related to habitual ways of viewing the world. In particular, the present finding relates to the more general problem of the existence of characteristic modes of perceiving which differ between different individuals. Some authors, notably Werner, have taken the position that physiognomic and geometric-technical qualities characterize dichotomous properties of perceptual objects and represent two distinct modes of perceiving. Werner, for example, states: "Art experience, we believe to be understood psychologically only if one discerns the two kinds of perception dubbed here 'physiognomic' and 'geometric-technical.' The building elements of art and aesthetic experience are physiognomic characters and physiognomic relations; the building elements of science, or realistic experience of objects, are geometric-technical (in their most abstract form: numerical) data" (9). The evidence from this experiment accords with this view in that such a conceptualization is reflected in the greater potency of physiognomic qualities in the artists' mode of cognition than in the chemists'.

In addition, the data obtained in this experiment also provide further substantiation of the general effect of directional dynamics on autokinetic and real motion. As in a previous experiment by Comalli, Werner, and

Wapner (1), it has been found that autokinetic motion predominantly occurs in the direction of the dynamics in a pictured object. With respect to real motion, here, as in a study by Wapner and Werner (6) it has been found that a moving pictured object with dynamics in the direction of motion appears to move faster than an object without directional dynamics.

REFERENCES

1. COMALLI, P. E., JR., WERNER, H., & WAPNER, S. Studies in physiognomic perception: III. Effect of directional dynamics and meaning-induced sets on autokinetic motions. *J. of Psychol.*, 1957, **43**, 289-299.
2. KADEN, S. E., WAPNER, S., & WERNER, H. Studies in physiognomic perception: II. Effect of directional dynamics of pictured objects and of words on the position of the apparent horizon. *J. of Psychol.*, 1955, **39**, 61-70.
3. KOFFKA, K. Principles of Gestalt Psychology. New York: Harcourt, 1935.
4. LANGFELD, H. S. The Aesthetic Attitude. New York: Harcourt, 1920.
5. MILLER, A., WERNER, H., & WAPNER, S. Studies in physiognomic perception: V. Effect of ascending and descending gliding tones on autokinetic motion. *J. of Psychol.*, 1958, **46**, 101-105.
6. WAPNER, S., & WERNER, H. Perceptual Development. Worcester, Mass.: Clark Univ. Press, 1957.
7. WERNER, H., & WAPNER, S. Studies in physiognomic perception: I. Effect of configurational dynamics and meaning induced sets on the position of the apparent median plane. *J. of Psychol.*, 1954, **38**, 51-65.
8. WERNER, H. Comparative Psychology of Mental Development. New York: International Univ. Press, 1957.
9. ———. On physiognomic perception. In (G. Kepes) *The New Landscape*. Chicago: Theobald, 1956.

Department of Psychology
Clark University
Worcester 10, Massachusetts

SLEEP SUGGESTIBILITY TEST*

Centro de Psiquiatria Experimental, Universidad de Chile

JULIO M. DITTBORN, O. GUTIERREZ, AND L. M. GODOY

A method has been devised which quantifies sleep suggestibility in subjects in whom adequate expectation has previously been inserted, and who, after being submitted to the body sway test, were considered to be good swayers (3). A résumé of this method has already been published (1).

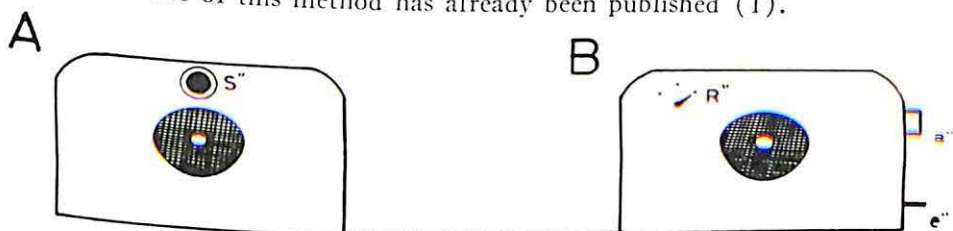


FIGURE 1

The test instrument consists fundamentally of an intercommunicator with two loudspeakers *A* and *B* (Figure 1). Loudspeaker *A* has a small neon bulb "S" (1/4 watt) mounted on its panel, which lights up dimly at regular intervals. These intervals are counted in seconds and their frequency regulated by a rotatory key "R" set in loudspeaker *B*. Loudspeaker *B* commands the whole intercommunication system by means of key "e"; it also has a special gadget "a" attached which gives the operator an auditive signal every time the bulb "S" lights up.

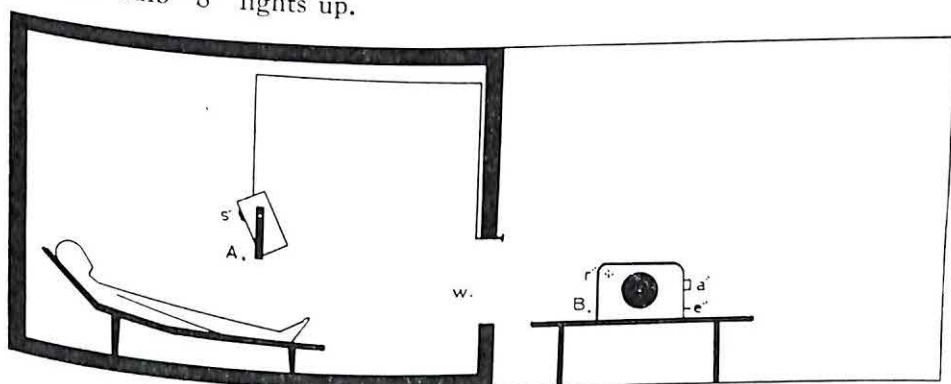


FIGURE 2

* Received in the Editorial Office on October 15, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

The *S*, who lies on a couch in a soundproof booth, is requested to repeat the word "dormir" (to sleep) every time the light signal appears on loudspeaker *A*. The operator hears the word through loudspeaker *B*, which is located in an adjoining room, while observing the *S* through window "*w*" (Figure 2). At a certain moment, the given *S*, in whom adequate expectation has been instilled, fails to respond to some of the light signals since his eyelids begin to close. Sleep suggestibility is measured by adding up the number of signals which have received no response in the individual cases. A *S* may sometimes skip one, two, or more signals and then continue to answer, while another may suddenly stop responding. We have never gone beyond the first 100 responded signals in any of our experiments.

TABLE 1
SLEEP SUGGESTIBILITY TEST TABULATION

Last signals receiving response	Signals receiving no response
15	1
21	3
22	18
	22 interruptions in 3 intervals

The test tabulation of a given experiment is illustrated in Table 1. In the right-hand column, the number of signals receiving no response is listed in separate groups after the last signal responded which, in turn, are listed in the left-hand column. For terminological purposes, a signal receiving no response has been called an "interruption," and an interruption or group of interruptions, an "interval." If a *S* fails to respond to 18 successive light signals (interruptions), we may say arbitrarily that he has attained a state of suggested sleep.

The state attained is denominated suggested sleep, and not natural sleep which can follow a monotonous action, since without adequate expectation (2) of the phenomenon, the susceptible *S* in a control experiment may continue to answer light signals far in excess of the first hundred perceived.

REFERENCES

1. DITTBORN, J. M., & KLINE, M. V. An instrument for the measurement of sleep induction. *J. of Psychol.*, 1958, **46**, 277-278.
 2. DITTBORN, J. M., & ARMENGOL, V. Expectation as a factor of sleep suggestibility II. *J. of Psychol.*, 1960, **49**, 113-116.
 3. FURNEAUX, W. D. Primary suggestibility and hypnotic susceptibility in a group situation. *J. Gen. Psychol.*, 1952, **46**, 87-91.
- Casilla (P.O. Box) 6507
Santiago, Chile
South America



EXPECTATION AS A FACTOR OF SLEEP SUGGESTIBILITY: II*

Centro de Psiquiatria Experimental, Universidad de Chile

JULIO M. DITTBORN AND VILMA ARMENGOL¹

A. PROBLEM

Expectation as a personality factor may be defined as an attitude implying a different degree of awareness towards a given situation which will subsequently result in a given type of behaviour. Expectation appears to be of importance in sleep suggestibility.²

We shall now describe a new set of experiments aimed at verifying this assumption that use a larger number of Ss in a different experimental situation.

B. SUBJECTS AND METHOD

Twelve healthy soldiers of the Chilean army that would be 20 years of age in 1958 were chosen as Ss because of their highly positive response to the body sway test. They were submitted to the sleep suggestibility test (2) in a soundproof booth with the operators (two in these experiments) in an adjoining room. The experiments were performed in the afternoon after a morning of heavy military duty.

Careful precautions were taken to prevent the Ss' becoming aware of our experimental intentions and to avoid communication among them while the experiments were being carried out. The Ss were given a standard questionnaire before and after each individual test either to check these points or to obtain further information.

Light signal frequency in the experiments varied from 7.5 to 11 seconds. No more than a hundred signals were made in any experiment.

In the first set of experiments, the Ss were requested to repeat the word "dormir" (to sleep) after each light signal.

In the second group of experiments, the following expectation was inserted in the Ss before each test started: "Now, you will feel more and more sleepy as long as you keep on repeating 'dormir' after each light signal."

* Received in the Editorial Office on October 15, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ Local Rockefeller fellow.

² Expectation as a factor of sleep suggestibility. *J. Clin. & Exp. Hypn.*, 1958, 6, 164-170, contained a number of printing errors; in Vol. 7, No. 3, erratum was published.

A good understanding of the expectation is of prime importance, so each *S* was asked to repeat the above-mentioned words before the pertinent experiment began.

A *S* was considered to be in a state of suggested sleep when he failed to respond to 18 successive light signals within the first 100 to be perceived.

Experiments with and without expectation were carried out on different days.

C. RESULTS

The results are shown in Table 1. With the exception of *S* No. 3, all the others had to receive adequate expectation before they could attain varying degrees of sleep suggestibility.

Some observations on experiments with *Ss* in whom expectation had been instilled are worth mentioning: (a) After some interruptions, two *Ss* failed to respond to a signal, although it was clear to the experimenter that their eyes were wide open. (b) Psychological disposition is important in sleep attainment. The degree of disposition was inferred after the pertinent questions had been posed; one *S* rapidly fell into suggested sleep, although he admitted he was trying very hard not to. If a given *S* persists in looking round the room between signals, it is often a sign of non-disposition. A few *Ss* said that they could have attained suggested sleep before the actual results were obtained if they had been told to let themselves sink into sleep instead of to answer the light signals as they assumed.

A pilot experiment was performed with the two loudspeakers, and the experimenter in a silent room with one *S* capable of prompt suggested sleep; to our surprise, the *S* kept repeating "dormir" for a while after his eyes had closed on hearing the click signal of the operator's loudspeaker only.

With the exception of one *S*, all awoke from suggested sleep after the operator's verbal intervention.

D. DISCUSSION

In order to attain a state of suggested sleep, it seems extremely important that a susceptible *S* should have a high degree of awareness of what he is supposed to do. We say a "high degree" because we consider that a certain degree of expectation has already been implanted by the experimental situation *per se*, in which the given *S* lies on a couch in a very dimly-lit room repeating the word "dormir" monotonously over and over again at regular intervals. Some *Ss*, in fact, reported a sleepy sensation, and the operator detected signs of this phenomenon, such as a lower voice tone, a tendency to sprawl more comfortably on the couch, a longer interval between signal and

TABLE 1
A COMPARISON OF RESULTS IN Ss WITH AND WITHOUT SLEEP EXPECTATION WHEN SUBMITTED TO THE CORRESPONDING TEST

Subjects	1	2	3	4	5	6	7	8	9	10	11	12
No expectation												
Interruptions	0	0	18	1	0	0	0	0	0	0	0	0
Intervals			1	1								
Expectation												
Interruptions	5	19	18	39	27	1	25	1	0	25	18	18
Intervals	5	2	1	10	9	1	6	1	0	5	1	1

answer, etc. The sensation was not, however, strong enough to make the S's eyes close within the span of the first 100 signals to be perceived, which, in our opinion, is a sign that the deepest state of sleepiness has been reached.

This state of suggested sleep has a positive electroencephalographic correlation, as briefly reported elsewhere (1).

With regard to psychological disposition, our Ss were chosen after being submitted to a body sway test on a specific afternoon in company with all their regimental comrades on the superior officer's command. A few were excluded from the experimental group because of illiteracy, and the remainder, that is, the group itself, were not questioned any further as to whether they really wished to co-operate in our experiments. In other words, no further investigation was made of their degree of willingness before the experiments began.

It seems worth while to mention that before the regiment as a whole was submitted to the body sway test, an address was delivered to the soldiers in which they were told very vaguely that those chosen were to collaborate with the University of Chile in experiments related to experimental psychology.

E. SUMMARY

1. A definition of expectation as a personality factor is given.
2. On this basis, the highly positive influence of expectation is demonstrated with respect to the entry into a state of quantified suggested sleep of a group of healthy men, who are about the same age and highly responsive to the body sway test.
3. The importance of psychological disposition, and what may be termed degrees of expectation and other factors such as willingness and the Ss' knowledge of the experimental situation, are discussed in relation to the results achieved.

REFERENCES

1. BORLONE, M., DITTBORN, J. M., & PALESTINI, M. Correlaciones electroencefalográficas dentro de una definición operacional de hipnosis. *Neurocirugia*, 1959, 17, 67-68.
2. DITTBORN, J. M., GUTIERREZ, O., & GODOY, L. M. Sleep suggestibility test. *J. of Psychol.*, 1960, 49, 111-112.

Casilla (P.O. Box) 6507
Santiago, Chile
South America

AN OPERATIONAL DEFINITION OF SOMNAMBULIST HYPNOSIS*

Centro de Psiquiatria Experimental, Universidad de Chile

JULIO M. DITTBORN AND VILMA ARMENGOL¹

A. PROBLEM

A method has been described which quantifies the entry into a state of suggested sleep in Ss that are highly responsive to the body sway test and in whom very definite expectations have been inserted (1, 2). With one exception, all Ss awoke from this kind of sleep after the operator's verbal intervention. This naturally led to the following question: what would be the behaviour of a given S if, as well as the expectation of sleep, the expectation of being able to speak without awakening were instilled in him?

B. SUBJECTS AND METHOD

Seven healthy young men of about the same age who were serving in the Chilean army in 1958 were used as Ss.² All had in common the capacity to attain suggested sleep, that is to say, they were unable to respond to 18 successive light signals within the first 100 to be perceived in the test given to them.

The experiment was performed while the S was lying on a couch in a soundproof room with the operator observing and controlling the test instruments from an adjoining room.

Apart from the expectation of sleep, the expectation of being able to speak without awakening was instilled into the individual Ss before the start of each experiment. A good understanding of the situation aimed at is important if clear results are to be obtained. It was sometimes necessary to illustrate the requisite expectation by citing the very common experiences of people who talk in their sleep.

Precautions were taken to prevent communication among the Ss.

After suggested sleep was attained, each S was asked a simple question. The questions called for a simple answer with meaning for the individual Ss,

* Received in the Editorial Office on October 15, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ Local Rockefeller fellow.

² These Ss belonged to the group of 12 on whom the experiments in reference (2) were made.

such as the Christian names of two friends or relatives, two numbers of their own choice between one and 10, etc. Two to three minutes after replying, the *S* was requested to open his eyes. Immediately after awakening, he was asked a generalized question in order to obtain his comments on what he thought the whole experience was about. If his reply did not evoke the sleep dialogue, a second question was put to him in the following standard words: tell me what happened while you had your eyes closed. It is our opinion that this makes no suggestion of amnesia, either directly or indirectly, and, on the contrary, may serve to stimulate the *S*'s memory. The degrees of the *S*s' evocation capacity were classified as follows: (a) total amnesia; (b) partial amnesia, i.e., the question is remembered but not the answer, or only one name or number given is recalled, etc.; (c) momentary amnesia: the *S* remembers the whole dialogue 30 seconds after the relevant question is posed; (d) no amnesia.

C. RESULTS

The results are given in Table 1. No. 1 was the only *S* who did not awake after the operator's verbal intervention, and who had not been told that he might expect to speak without awakening. In fact, he emphatically denied that he had been informed by anyone of the possibility of such behaviour.

TABLE 1
Ss CAPACITY TO DEVELOP DIFFERENT DEGREES OF SPONTANEOUS AMNESIA AS DEFINED WITH RESPECT TO A SIMPLE AND MEANINGFUL CONTENT AFTER SUGGESTED SLEEP HAS BEEN ATTAINED

Subjects	1	2	3	4	5	6	7
Expectation of sleeping and of speaking without awakening							
Interruptions	19 ^a	19	31	18	18	18	21
No amnesia	2	2	6	1	1	1	3
Partial amnesia				x			
Momentary amnesia	x				x		
Total amnesia		x					
			x			x	x

^a This *S* could speak while in suggested sleep without the pertinent expectation.

Within the experimental framework, some observations are of particular interest: (a) When requested to sleep and to speak without awakening, some *S*s appear to fall into a deeper sleep, to judge by their type of respiration and occasional snoring. This contrasts with the previous experiments, when the expectation of sleep only was inculcated; (b) the pronunciation of most *S*s was less clear in the somnambulistic than in the waking state; (c) when

asked an equivalent question, the Ss took longer to reply when in a state of suggested sleep than when awake.

D. DISCUSSION

One of the main drawbacks to hypnosis as a method of clinical and experimental investigation is, perhaps, the vagueness of its limits as a psychophysical entity. Vagueness in circumscribing the hypnotic phenomenon leads, *inter alia*, to the irreproducibility of experiments performed within the hypnotic state. In our opinion, such a situation greatly impedes the possibility of substantial advances in this field.

Hypnotism is a term that covers a wide range of psychophysical signs, most of which have in common the fundamental feature of being induced artificially by suggestion. If we observe closely a typical case of hypnotic induction, we can see that the operator mainly mobilizes his subject's suggestibility at the level of two biological functions; on one level, the suggestions act predominantly on the voluntary motor system (body sway, catalepsy of eyes and limbs, muscular relaxation, etc.) while the other reception level is the sleep function. Suggestions to sleep are given directly, indirectly, or in a marginal way, i.e., by means of a neutral proposition made in a calm and drowsy tone of voice. This method of inducing the hypnotic state is in line with what we know today of the high correlation between primary suggestibility, which mainly concerns the positive responses of the motor system, and the phenomenon of hypnosis. The more a subject is responsive to a primary suggestibility test, the greater is his possibility of passing into a deep hypnotic trance.

Deep hypnotic trance is also known as the somnambulistic state of hypnosis. If we consider this state as the most constant of, and phenomenologically richer than, the states of hypersuggestibility that are related to functions other than that of sleep, we may define the hypnotic situation as follows: "A psychophysical entity induced by direct, indirect, or marginal suggestions to sleep, closely resembling natural somnambulism and characterized by spontaneous and induced modifications of biological functions which differ both quantitatively and qualitatively from those that can be found or induced in the so-called waking or normal state." This definition excludes hypersuggestibility related to a psychophysical function operating outside a state of suggested sleep. We already know that within the somnambulistic state the suggestibility of the most diverse functions is greatly enhanced.

This general definition indicates the possibility of giving the somnambulistic state an operable instrumentation. We therefore developed the sleep sug-

gestibility test (1), and discovered that expectation was necessary for the attainment of suggested sleep within our experimental situation (2) and that, by means of further expectation, a state of verbal communication could be achieved with susceptible subjects who had different degrees of spontaneous amnesia with regard to a meaningful content for a short period after the operator's verbal intervention. Hence, we propose the following operational definition of somnambulistic hypnosis: "The capacity of a given subject, in whom adequate expectation has been inserted, to enter into a state of suggested sleep, that is, to be unable to respond to 18 consecutive light signals out of the first 100 to be perceived in the corresponding test, and to maintain verbal communication with his operator, this capacity being accompanied by different degrees of spontaneous amnesia, as defined, with respect to a simple and meaningful content for at least two minutes after the verbal contact."

This operational definition limits the hypnotic phenomenon to the somnambulistic state and this state to a mere verbal contact with the experimenter. However, we consider that it will give us a frame of reference for future clinical and experimental communications. Another advantage of the situation is that the experimenters' personal qualifications play little or no part in the production of the hypnotic phenomenon; in our laboratory, many experimenters with a minimum of training are able to produce and measure the somnambulistic state, as defined, with no significant differences among them so far.

In clinical practice, we have been inducing somnambulism by requesting a prospective subject, who is in an adequate state of expectation, to whisper the word "duermo" (I sleep) after a complete respiratory movement. This technique enables us to observe that a susceptible subject lowers his voice tone at a certain moment and thereafter ceases to whisper the given word. The typical relaxation of the motor system that accompanies the sleep state can be easily perceived at that moment. Sleep suggestibility is tabulated by this method, respiratory movements being used as points of reference instead of the light signals. The verification of the different degrees of amnesia is controlled as described.

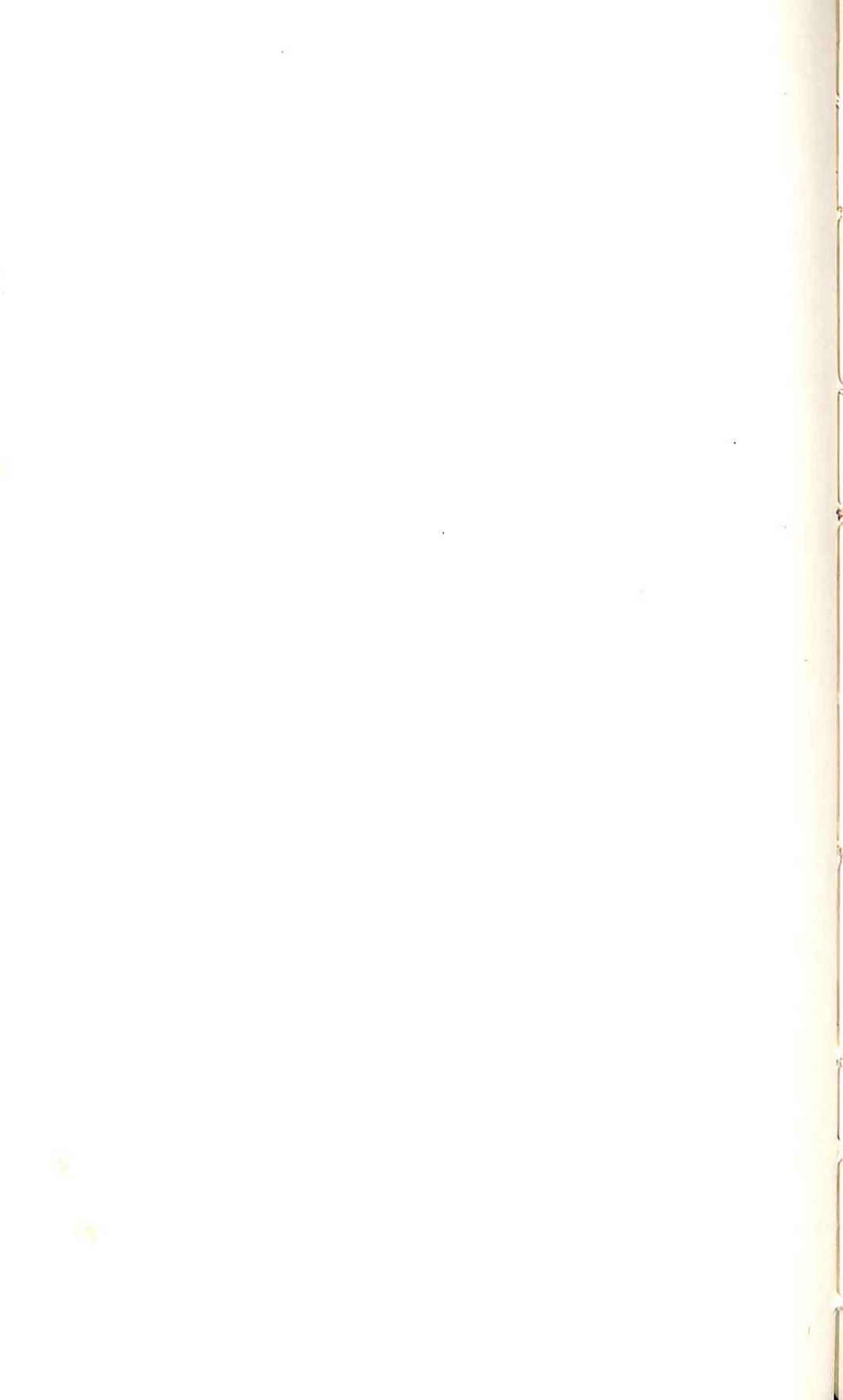
E. SUMMARY

1. Subjects susceptible to suggested sleep need to have the pertinent expectation of being able to speak while asleep without awakening.
2. Somnambulism is defined as this capacity, plus a different degree of spontaneous amnesia with respect to the verbal contact after a given time.
3. A discussion follows which gives an operational definition of somnambulistic hypnosis and is intended to serve as a future frame of reference for clinical and experimental communications.

REFERENCES

1. DITTBORN, J. M., GUTIERREZ, O., & GODOY, L. M. Sleep suggestibility test. *J. of Psychol.*, 1960, **49**, 111-112.
2. DITTBORN, J. M., & ARMENGOL, V. Expectation as a factor of sleep suggestibility:
II. *J. of Psychol.*, 1960, **49**, 113-116.

Casilla (P.O. Box) 6507
Santiago, Chile
South America



JOB COUNSELING OF THE EPILEPTIC*

Office of Psychological Services, Fordham University

WILLIAM P. ANGERS

A. INTRODUCTION

In the counseling situation the counselor, besides drawing from his training and experience, is presented with a different kind of challenge when the counselee is physically and/or emotionally handicapped. This requires a special kind of knowledge and skill on the part of the counselor as well as a sensitivity to the particular problem on hand. Such is the challenge which the epileptic, among others, presents.

The purpose of this paper is twofold: One is to study the job opportunities available to the epileptic; and, two is to gain whatever insights are possible about the epileptic who comes for job counseling in order to render him better service because of a clearer understanding of vocational rehabilitation counseling of the epileptic.

B. DESCRIPTION OF THE EPILEPTICS

In a report¹ from the Department of Health, Education, and Welfare, Office of Vocational Rehabilitation, Washington 25, D.C., there was summary information available on 1,521 epileptics whose cases were closed as rehabilitated during the fiscal year 1957 under the State-Federal vocational rehabilitation program.

1. *General Characteristics*²

Among this number there were 1,077 or 70.8 per cent who were men although the incidence of epilepsy is about the same for both sexes (16).

In regards to race there were 1,250 or 90.9 per cent who were white. While seizures occur among all races, "exact comparable statistics" are not available according to W. G. Lennox (17).

* Received in the Editorial Office on October 30, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ Rehabilitation Service Series Number 450—Supplement 6. *Facts in brief—persons with epilepsy rehabilitated in fiscal year 1957*, July, 1958.

² Cases for whom data were not reported were excluded from the total when computing per cents, averages, means, and medians for sex, race, marital status, age at acceptance, age of onset, years of schooling, origin of disability, and estimated annual earnings.

2. *Marital Status*

There were 953 or 62.7 per cent who were never married. Perhaps more would have been married if it were not prohibited or forbidden in some 17 states (20).

The average age of the 1,521 epileptics at the time of their acceptance for vocational rehabilitation counseling was 25 years whereas 893 or 59.2 per cent were under 15 at the age of onset, 112 or 7.4 per cent who were from 30 to 44, while 32 or 2.1 per cent were 45 years or older at the age of onset. There were 12 whose age was not reported. Thirteen was the median age at disablement. These figures are in keeping with the findings of D. T. Davidson (7) concerning the age of onset for the occurrence of epilepsy.

The average school time for this group of epileptics was 10 years although more may be expected of them in terms of their education (9, 14).

3. *Origin of Disability*

For 1,196 or 78.6 per cent "disease" was the cause of their disability; 179 or 11.8 per cent it was "congenital"; 134 or 8.9 per cent it was "accidents"; 11 or 0.7 per cent it was military or naval service; while for one it was not reported. These figures indicate how little is known about the cause of this illness [F. A. Gibbs (10); W. G. Lennox (16); W. Penfield and K. Kristiansen (21); W. Penfield and H. Jasper (22); T. J. Putnam (23); H. Yahraes (24)] and surely confirm the need for the encouragement and support of the various organizations which are working with epileptics and which are informing the public.³

The referral sources were as follows: 348 or 22.9 per cent from doctors, hospitals, etc.; 182 or 12.0 per cent from public welfare agencies; 46 or 3.0 per cent from other welfare agencies; 224 or 14.7 per cent from the State Employment Service offices; 227 or 14.9 per cent from schools; 206 or 13.6 per cent from interested persons; 155 or 10.2 per cent were self-referred; and 133 or 8.7 per cent were from public officials, employers, labor unions, etc.

4. *Time on Rolls*

The median from the time of the referral to acceptance for the 1,521 epileptics was three months with a mean of five months, whereas from the

³ Most large cities in the U.S.A. have epilepsy agencies, centers, or leagues. A few such organizations to whom this writer is grateful for their help: Epilepsy Information Center, 73 Tremont Street, Boston, Mass.; National Epilepsy League, Inc., 208 North Wells Street, Rooms 201-202, Chicago 6, Ill.; and the United Epilepsy Association, 113 West 57th Street, New York 19, N.Y. There is also the U.S. Office of Vocational Rehabilitation, Department of Health, Education, and Welfare, Washington 25, D.C., and the regional offices throughout the country.

time of acceptance to closure there was a median of 14 months with a mean of 17 months. This gives an indication of the time expectancy to be used in the satisfactory job counseling of the epileptic.

C. RESULTS AND DISCUSSION

In addition to guidance, counseling, and placement, the following services were provided: 453 or 29.7 per cent received training; 186 or 12.3 per cent received training and physical restoration. Of these numbers 601 or 39.5 per cent received vocational training. There were 307 or 20.2 per cent who received physical restoration; and 26 or 1.7 per cent who received other services. Among these numbers there were 549 or 36.1 per cent who received guidance, counseling, and placement only.

The results of the 601 or 39.5 per cent who received vocational training are presented in Table 1. It is interesting to note the large number who went to business school, college, or university. These results suggest that when there is no obstacle placed in the path of the epileptic by the school as found in a study by Lennox, McBride, and Potter (18), that vocational training is not only appropriate but advisable when the job market is taken into consideration with its present-day requirements.

TABLE 1
A BREAKDOWN OF THE 601 OR 39.5% OF THE 1,521 EPILEPTICS RECEIVING VOCATIONAL TRAINING

Type of training	Number	Per cents
Business college	178	29.6
Private trade school	99	16.5
Public vocational school	84	14.0
Sheltered workshop	29	4.8
College or university	101	16.8
Business establishment	70	11.6
Tutor	18	3.0
Correspondence school	4	0.7
Other	18	3.0

A complete list of the jobs held by the 1,521 epileptics who were considered rehabilitated at closure, fiscal year ending June 30, 1957, is presented in Table 2. The jobs were listed according to the guide which is provided by the *Dictionary of Occupational Titles* (11).

When it is considered that the additional jobs were included which resulted from the vocational training, it indicates that the job opportunities which are available to the epileptic are about as numerous as to any others providing that there is no handicap either by the epileptic (depending on

TABLE 2
NUMBER OF PERSONS WITH EPILEPSY REHABILITATED, BY JOB OR OCCUPATION
AT CLOSURE, FISCAL YEAR ENDED JUNE 30, 1957

Job or occupation at closure	Number	Job or occupation at closure	Number
Professional occupations, total.....	38	Clerical & kindred occu.—Con.	
Accountants and auditors	7	Secretaries	16
Chemists	1	Shipping & receiving clerks	14
Clergymen	5	Statistical clerks & compilers	4
Engineers, civil	3	Stenographers and typists	36
Engineers, electrical	2	Stock clerks	61
Musicians & teachers of music.....	4	Telephone operators	2
Pharmacists	2	Other	2
Teachers (secondary school) and principals	6	Sales & kindred occupations, total	77
Teachers & instructors, schools for handicapped, vocational, etc.	1	Canvassers and solicitors	5
Veterinarians	2	Salesmen, insurance	2
Natural scientists	3	Newsboys	2
Other	2	Hucksters and peddlers	6
Semiprofessional occupations, total	36	Sales clerks	37
Decorators and window dressers..	2	Salespersons	10
Commercial artists	2	Salesmen, to consumers	13
Draftsmen	14	Salesmen and sales agents, except to consumers	1
Laboratory technicians and assts.	9	Sales clerks, dry cleaning and laundry	1
Photographers	3	Service occupations, total	268
Surveyors	1	Domestic service, total	59
Embalmers and undertakers	2	Dayworkers	6
Technicians, except laboratory	3	Laundresses, private family	1
Managerial and official occupations, total	15	Housekeepers, private family	8
Hotel and restaurant managers....	2	Housemen and yardmen	9
Retail managers	6	Maids, general	17
Inspectors, public service	4	Nursemaids	12
Other	3	Parlormaids	2
Clerical and kindred occupations, total	312	Miscellaneous servants, private family	4
Bookkeepers and cashiers, except bank cashiers	42	Personal service, total	141
Bookkeeping machine operators....	2	Bartenders	3
Checkers	6	Boarding-house and lodging-house keepers	2
Clerks, general	10	Maids and housemen, hotels, restaurants, etc.	4
Clerks, general office	61	Cooks, except private family.....	4
Financial institution clerks	1	Waiters and waitresses, except private family	19
Printing and publishing clerks....	3	Kitchen workers in hotels, restaurants, railroads, etc.	49
Clerks in trade	3	Barbers, beauticians, manicurists	13
File clerks	8	Practical nurses & nurses aides....	16
General industry clerks	20	Attendants, recreation and amusement	5
Messengers, errand boys, and office boys and girls	8	Attendants, hospitals and other institutions	23
Office machine operators	10	Attendants, professional and personal service	1
Paymasters, pay-roll clerks, and timekeepers	1		
Post office clerks	1		
Mail carriers	1		

Note.—Date partially estimated.

TABLE 2 (Continued)

Job or occupation at closure	Number	Job or occupation at closure	Number
Service occupations—Con.		Skilled occupations—Con.	
Personal service—Con.		Welders and flame cutters	3
Doormen	1	Occu. in fabrication of metal products	1
Ushers	1	Occu. in mfg. of electrical machinery and accessories	14
Protective service, total	19	Occu. in manufacture of automobiles	2
Guards and watchmen, except crossing watchmen	10	Occu. in mfg. fabricated of plastics products	2
Soldiers, sailors, marines, etc.....	9	Painters, except construction and maintenance	5
Building service, total	49	Miners and mining machinery operators	2
Charwomen and cleaners	1	Brick and stone masons and tile setters	2
Janitors and sextons	34	Carpenters	3
Porters	12	Cement and concrete finishers.....	2
Elevator operators	2	Painters, construction and maintenance	4
Agricultural and kindred occupations, total	116	Plumbers, gas fitters, and steam fitters	1
General farmers	28	Locomotive firemen	2
Animal and livestock farmers	4	Linemen and servicemen, telegraph, telephone, power	2
Poultry farmers	7	Occu. in laundering, cleaning, dyeing, pressing, etc.	1
Farm hands, grain, dairy, general farms, animal & livestock, etc.	56	Meatcutters, except in slaughtering and packing houses	4
Fruit & vegetable graders and packers	1	Engineers, stationary	1
Nursery and landscaping laborers ..	7	Millwrights	2
Gardeners and grounds keepers, parks, cemeteries, etc.	12	Mechanics and repairmen, motor vehicle	24
Fishermen and oystermen	1	Mechanics and repairmen, radio, TV, etc.	23
Skilled occupations, total	163	Photographic process occupations ..	2
Bakers	1	Foremen	3
Occu. in manufacture of knit goods	2	Semiskilled occupations, total	228
Furriers and occu. in fabrication of fur goods	1	Occu. in production of bakery products	2
Dressmakers & seamstresses	7	Occu. in processing of dairy products	2
Cabinetmakers	4	Occu. in manufacture of knit goods	3
Upholsterers	11	Nonprocess occupations in mfg. of textiles	2
Occu. in mfg. of paper goods	2	Other occu. in mfg. of textiles.....	3
Compositors and typesetters	1	Dressmakers & seamstresses	4
Pressmen and plate printers, printing	3	Occu. in fabrication of textile products	5
Other occu. in printing & pub.....	2	General woodworking occu.	6
Occu. in production of industrial chemicals & chemical products	3	Upholsterers	1
Shoemakers and shoe repairmen, not in factory	3		
Jewelers, watchmakers, goldsmiths, & silversmiths	8		
Machinists	2		
Tinsmiths, coppersmiths, and sheet metal workers	4		
Molders	2		

TABLE 2 (Continued)

Job or occupation at closure	Number	Job or occupation at closure	Number
Semiskilled occupations—Con.		Semiskilled occupations—Con.	
Occu. in mfg. of furniture	1	Linemen and servicemen, tele- phone, telegraph, and power	1
Occu. in mfg. of miscellaneous finished lumber products	2	Motion picture projectionists	3
Occu. in printing & publishing.....	7	Occu. in laundering, cleaning, dyeing, etc.	5
Occu. in production of plastics and related synthetic materials, and in the production of chem- icals	3	Attendants, filling station and parking lots	27
Occu. in refining of petroleum	2	Packing, filling, labeling, mark- ing, bottling, washing, and re- lated occupations	7
Occu. in mfg. of boots & shoes.....	2	Firemen other than process fire- men	2
Machine shop and related occu.	5	Oilers of machinery	1
Occu. in fabrication of metal products	2	Inspectors	4
Inter-industry metalworking occu.	1	Mechanics & repairmen, airplane	1
Occu. in building of aircraft	2	Mechanics and repairmen, motor vehicle	7
Semiskilled workshop operators....	14	Mechanics and repairmen, other	9
Occu. in manufacture of fabrica- ted plastics products	4	Transportation equipment la- borers, washers and greasers....	4
Occu. in the manufacture of mis- cellaneous products	4	Photographic process occu.	4
Painters, except construction and maintenance	5	Chainmen, rodmen, and axmen, surveying	5
Pattern and model makers, except paper	1	Warehousing, storekeeping, hand- ling, loading, unloading, and related occupations	12
Miners, mining machine operators and occupations in extraction of minerals	2	Carpenters' apprentices	3
Brick and stone masons and tile setters	2	Plumbers' apprentices	2
Painters, construction and main- tenance	3	Apprentices to other construction and hand trades	1
Roofers and slaters	2	Apprentices to printing trades....	3
Construction occupations, other....	12	Apprentices to other trades	7
Routemen	1	Miscellaneous occupations	146
Chauffeurs and drivers, bus, taxi, truck, and tractor	10	Unskilled occupations, total	146
Other transportation occu.	4	Laborers	40
		Family workers	82
		Housewives	1,521
		TOTAL	

how he feels about his disability) or by the employer (depending on how he feels about the disability).

The experience of this psychologist in job counseling of epileptics and for which one must be on guard is their compensation for their problem(s) by choosing unsuitable occupations. For instance, there were two who were plasterers and reported injurious accidents when an attack came while they were on a ladder, while another epileptic was a salesman who had difficulty in meeting and getting along with people so that he often had "fits" when

trying to sell a product to a customer. In these few instances each one was assisted in changing to a more suitable occupation in which the individual felt adequate, and a better social and emotional adjustment was reported. Other examples could be cited but the relevancy would be only additive.

The dynamics of compensation for disabilities was sounded by the pioneer work of Alfred Adler (2). He developed this concept to its psychological conclusion in a positive beneficial way to the individual and mankind in a later work (3). A study of the effects of negative compensation on workers in industry was carried out by his daughter, Alexandra Adler (1), which is also very useful in understanding this problem.

Another point worthy of consideration by the counselor is personal adjustment counseling of the epileptic which is often requested. It has been found by this writer that the amount of counseling should be proportional to the degree of prejudice to which the epileptic has been exposed.

In addition to the many cues which may be obtained from Table 2, there is also another excellent source of information which may be used as a guide for the counselors in their use of psychological tests with the epileptic. A detailed analysis of worker characteristics as they relate to the tests results are presented in tabular form in the *Estimates of Worker Trait Requirements For 4,000 Jobs as Defined in the Dictionary of Occupational Titles* (13).

The result of job groupings is presented in Table 3. It is striking how many are employed in skilled, semiskilled, clerical, sales, and kindred jobs

TABLE 3
A BREAKDOWN OF THE MAJOR OCCUPATIONAL GROUPS AFTER REHABILITATION
OF THE 1,521 EPILEPTICS

Occupational category	Number	Per cent
Skilled and semiskilled workers	391	25.7
Clerical, sales, and kindred jobs	389	25.6
Service workers	268	17.6
Unskilled workers	146	9.6
Family workers and homemakers	122	8.0
Agricultural and kindred workers	116	7.6
Professional, semiprofessional, and Managerial occupations	89	5.9

Note.—91 or 6.0 per cent of the 1,521 persons rehabilitated were self-employed.

which may be due to the ease by which these jobs are obtained, often with little or no information about the applicant and it may also point out the difficulty which the epileptic faces in obtaining employment. Both of these findings are in agreement with the study of Brinn and Smith (5).

Of special interest to counseling psychologists is the economic status of

the epileptics at the time of their acceptance and at the time of their closure as rehabilitated.⁴ There were 1,299 or 85.4 per cent who were unemployed at acceptance. Of this number 274 or 18.0 per cent had never worked before. There were 973 or 64.0 per cent who were dependent on their families for major support; 159 or 10.5 per cent on relief for major support (of this number 149 or 9.8 per cent were on public relief); 428 or 28.1 per cent were married, and 466 or 30.6 per cent had dependents.

There were 161 or 10.6 per cent who were farmers or family workers and 1,360 or 89.4 per cent receiving earnings. The estimated annual earnings of the 1,360 persons the first year after rehabilitation was \$3,128,100. There was an increase of earnings of the group from \$279,300 *before* rehabilitation to \$3,128,100 *after* rehabilitation. These results strongly support the advisability of guidance, counseling, training, and placement of the epileptic which has already been recommended (8, 12, 15).

D. RESEARCH WITH THE EPILEPTIC

It would have been interesting to have done an *IQ* and subtest scores pattern analysis had the Wechsler-Bellevue Intelligence Test been assigned to this group of non-institutionalized epileptics,⁵ although Collins (6) found the average *IQ* of 400 office patients to be above the norm. Another interesting facet would have been a history of epilepsy in the family and whether this would have been in keeping with the findings of Lennox (19).

A job follow-up of five years may have yielded significant information in regards to the number of jobs held, the reasons for leaving employment, the time jobs were held, and a listing of occupations by code number classifications for maximal job differentiations.

Finally, the type of seizures, the aura, the *EEG* examination results, the medication, the amount and the frequency of the dosages may have proved worthy of further investigation.

All of these points indicate the direction which future studies on epilepsy might take as they relate to job counseling of the epileptic.

In order to assist counselors and psychologists in this work, the pertinent information described in this study should be gathered on a supplementary form for his personal use.

E. SUMMARY AND CONCLUSIONS

In terms of the twofold purpose of this study we have seen that there are any number of opportunities available to the epileptic but he may find

⁴ Note data partially estimated.

⁵ This writer (4) did such a study with institutionalized epileptics.

difficulty in obtaining employment either because of his physical and/or emotional handicap(s) or because of the attitude of the public towards this disability.

The information that was gathered concerning the epileptic from this study may be summarized as follows:

1. The number of men were appreciably more than the number of women although the incidence of epilepsy occurs equally among both sexes.
2. The largest number of these epileptics were of the white race although "exact comparable statistics" are not available.
3. There were a large number of epileptics who were not married which may be due to a restriction by some 17 states.
4. The average age of onset for the 1,521 epileptics was in keeping with the findings of other studies.
5. There was little that was known about the cause of their epilepsy. This bears out the necessity for the encouragement and support of the societies, associations, and organizations which not only assist these people, but also serve the vital rôle of informing the public about the facts of this disability so that a healthy, realistic public attitude may be taken towards the epileptic.
6. The effectiveness of job counseling, guidance, training, and placement of the epileptic was analyzed and discussed with cues for counselors and psychologists.
7. We saw the importance of considering compensation used by the epileptic in the selection of an occupation.
8. We saw that the amount of counseling should be proportional to the degree of prejudice to which the epileptic is exposed.
9. The information gained from this study, if representative of the epileptic, should prove helpful in the counseling process.
10. The direction which future studies of epilepsy should take was indicated.

REFERENCES

1. ADLER, ALEXANDRA. The Psychology of repeated accidents in industry. *Amer. J. Psychiat.*, 1941, **98**, 99-101.
2. ADLER, ALFRED. A Study of organ inferiority and its physical compensation. *Nerv. & Ment. Dis. Monog. Ser.*, 1917, No. 24.
3. ———. Social Interest: A Challenge to Mankind. London: Faber & Faber, 1938.
4. ANGERS, W. P. A psychometric study of institutionalized epileptics on the Wechsler-Bellevue. *J. Gen. Psychol.*, 1958, **58**, 225-247.
5. BRINN, C., & SMITH, E. E. Opportunities Limited. San Francisco, Calif.: California Society For Crippled Children, 1951.

6. COLLINS, A. L. Epileptic intelligence. *J. Consult. Psychol.*, 1951, **15**, 393-399.
7. DAVIDSON, D. T. Seizures and Epilepsy. Boston: Childrens Medical Center, 1954.
8. Epilepsy Information Center: The Person With Epilepsy at Work. Boston: EIC, 1957.
9. Epilepsy Information Center: Epilepsy and the School Child. Boston: EIC.
10. GIBBS, F. A. A Modern View of Epilepsy. Chicago: National Epilepsy League.
11. Government Printing Office. Dictionary of Occupational Titles. (Rev. ed.) Washington, D.C.: GPO, 1949. (Vols. 1-11.)
12. ———. Interviewing Guides for Specific Disabilities: Epilepsy Washington, D.C.: GPO, BES-USES, 1954.
13. ———. Estimates of Worker Trait Requirements For 4,000 Jobs as Defined in the Dictionary of Occupational Titles. Washington, D.C.: GPO, BES-USES, 1956.
14. ———. The Child with Epilepsy. Washington, D.C.: GPO.
15. LENNOX, W. G. Employment of epileptics. *Indus. Med.*, 1942, **11**, 571-574.
16. ———. Seizure States. In (J. M. Hunt, Ed.) *Personality and the Behavior Disorders* (Vol. 2). New York: Ronald, 1944. (Pp. 938-967.)
17. ———. Science and Seizures. (Rev. ed.) New York: Harper, 1946.
18. LENNOX, W. G., McBRIDE, M., & POTTER, G. The Higher education of epileptics. *Epilepsia*, 1947, **3**, 182-197.
19. LENNOX, W. G. The Heredity of epilepsy as told by relatives and twins. *J. Amer. Med. Assoc.*, 1951, **146**, 529-536.
20. LENNOX, W. G., & MARKHAM, C. H. The Sociopsychological treatment of epilepsy. *J. Amer. Med. Assoc.*, 1953, **152**, 1690-1694.
21. PENFIELD, W., & KRISTIANSEN, K. Epileptic Seizure Patterns. Springfield, Ill.: Thomas, 1951.
22. PENFIELD, W., & JASPER, H. Epilepsy and the Functional Anatomy of the Human Brain. Boston, Mass.: Little Brown, 1954.
23. PUTNAM, T. J. Convulsive Seizures. Philadelphia: Lippincott, 1943.
24. YAHRAES, H. Epilepsy—The Ghost is out of the Closet. New York: Public Affairs Pamphlets, 1954.

235 East 22nd Street
New York 10, New York

A COMPARISON OF ADJECTIVAL ANTONYMS BY SIMPLE CARD-PATTERN FORMATION*

Departments of Psychology, Rutgers University and University of Portland

BRUCE M. ROSS AND NISSIM LEVY

A. THE PROBLEM

Of interest to a description of the thinker as a consumer of categories (2) is the comparison of category labels that are commonly classed as antonyms. Osgood (3, pp. 152-153; 327-328), in his conceptualization of a semantic space and his semantic differential procedure which makes use of adjectival antonym pairs, has called attention to the desirability of measuring adjectival opposition by methods which are not dependent on the circular procedure of rating adjectives against other adjectives. The present study avoids rating scale difficulties in comparing antonyms, but is restricted to a simple production task, card-pattern formation elicited from three antonym pairs—beautiful-ugly, simple-complex, common-unusual.

It was hypothesized, for each of the three adjective pairs employed, that one of the adjectives in each antonym pair would produce a more compact distribution than the other—hence a less ambiguous use of that adjective. Our hypothesis was based on the peculiarities of customary usage which designate in esthetics (and possibly morals) "the good, the true, and the beautiful" as rare instances in any sample of events, if not entirely singular. This hypothesis of relative antonym differences is contrary to the assumption of the semantic differential. Also, in an attempt to locate possible developmental usage differences, card-pattern distributions of a fourth-grade group were compared with distributions from an Ivy League college student group.

B. METHOD

The problem was administered in elementary psychology laboratory classes at Brown University to obtain an adult sample of 100 Ss and to 100 fourth-grade students in classes from successive years in the Bristol, Rhode Island, public schools to obtain 100 child Ss. Not more than three faulty patterns were produced by either group for any adjective. Sex distribution for both groups was about equal.

Individual Ss were given 10 playing cards, five with red and five with

* Received in the Editorial Office on November 1, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

blue backs, and told to select any nine cards to form the "most beautiful pattern you can make by putting the cards in a square with three rows and three columns." Subjects marked the patterns they composed on previously prepared answer sheets. Subjects then shuffled all 10 cards and were asked to make and record the "ugliest" pattern possible in a three row by three column format. The procedure was repeated in the same way for pattern formation of "simplest," "most complex," "commonest," and "most unusual" in consecutive order. The same directions and procedures were adopted *mutatis mutandis* for the fourth-grade students.

An empirical classification system was established by the *Es* in classifying all obtained patterns. The large number of potentially different patterns produced were reduced to 23 pattern categories by collapsing the complete pattern classification in three ways. First, all color distinctions were disregarded as to whether red or blue was in the majority or minority. Second, all patterns that were alike if rotated were classed together. Third, mirror images of a given pattern and all rotations of the mirror image were likewise classed as a single pattern. Two patterns, the diamond and cross, had no rotational or mirror-image permutations. Eight other patterns had rotational permutations but no unique mirror-image permutations. The remaining 13 pattern categories had both unique rotational and mirror-image permutations. Thus, each of the pattern categories in this latter group had a total of eight possible permutations, or 16 permutations if color differences were tallied.

C. RESULTS AND CONCLUSIONS

The degree to which total frequency distributions of patterns produced to each antonym are compact or spread out can be indicated by expressing the amount of uncertainty in bits, the H statistic. (H is a biased estimate of the true value H , but Attneave (1, pp. 65-66) suggests that if many zero frequencies are present the usual correction formula cannot be appropriately applied; therefore all computed H values shown are uncorrected.) Table 1 presents the H values obtained for each adjective on a 23-category basis together with H_{rel} , an index varying from 0 to 1.00 that is the ratio of uncertainty in bits empirically produced to the maximum number theoretically possible ($\bar{H}/H_{max}$). The results in Table 1 show similarities between children and adults for "ugly," "complex," and "unusual"—high H_{rel} indices indicate many different patterns were formed with little favoritism for individual patterns. For "beautiful," favoritism resulted in a lower H_{rel} for both children and adults than for "ugly." It is with "simple" and "common" that the only major differences between children and adults

TABLE 1
SUMMARY OF RESULTS DERIVED FROM TOTAL DISTRIBUTIONS OF CHILD AND ADULT CARD-PATTERN FORMATION

Statistic	Beautiful		Ugly		Simple		Complex		Common		Unusual	
	C	A	C	A	C	A	C	A	C	A	C	A
N	98	100	98	99	100	100	100	100	97	100	97	100
Number of categories used	17	13	22	20	20	12	21	21	19	6	22	23
H	2.29	2.40	3.86	3.85	3.79	2.85	3.90	3.80	3.31	1.74	4.08	4.15
H _{rel}	.51	.53	.85	.85	.84	.63	.86	.84	.73	.38	.90	.92

C = Children, A = Adults, $H_{max} = 4.52$.

Note—All calculations performed on basis of N = 100 per group and originally computed to four decimal places.

appear. For both of these adjectives the adults produced more compact distributions, as shown by lower H_{rel} values.

Although tests of significance are not available for the intercomparisons of different distributions, some interest attaches to the consistency with which values of approximately .85 appear as H_{rel} values indicating well dispersed distributions and consequent lack of category favoritism. This suggests that the *Es'* empirical category reduction scheme did not artificially produce category favorites to any large extent.

A second type of analysis is presented in Table 2 in which the frequency of repeats by the same *S* of the same pattern category for two different adjectives is shown; all repeat frequencies which are above seven for either child or adult groups are presented. The same 23-category classification scheme is used as with Table 1. The fact that H_{rel} values below the modal .85 level are found with some adjectives can be explained in part by this favoritism tendency to use the same pattern with two or more adjectives. High repetition frequencies further suggest that within the limits of the experimental situation partially overlapping concepts are being invoked. It is evident that with the adult sample "beautiful-common" and "common-simple" have stronger tendencies to produce repetitions than is the case with the children, while "beautiful-complex" is relatively favored as a repetition by the children. The former result implies the smaller uncertainty previously noted in the "common" distribution; while the latter result partly explains why the children, who used more categories than the adults for "beautiful," did not at the same time have greater distribution uncertainty.

TABLE 2
FREQUENCIES OF REPEATS OF THE SAME PATTERN CATEGORY BY THE SAME SUBJECT

Adjectives	Adults	Children
Beautiful-Common	39	18
Common-Simple	20	5
Beautiful-Simple	8	11
Beautiful-Complex	5	14

With reference to specific patterns produced, by far the most frequent contributor to repetitions with both child and adult samples was the completely symmetrical diamond pattern. The reduction of uncertainty with the "simple" distribution seemed to have an important source for the adults in a concept of homogeneous grouping of common elements without alternation; 34 adults used a single category of this sort as compared to 13 children. Except for this same pattern category, which was also produced by eight adult *Ss* to "common," all other "common" patterns had either vertical

or horizontal mirror-image symmetry. Some exploratory work reported by Ridderheim and Levy (4) bears on the finding that the children's distribution for "common" did not show a similar compactness by demonstrating that, as tested by a reversal learning situation, most children at the fourth-grade level have only partially mastered the bilateral symmetry concept.

We would conclude from our results that:

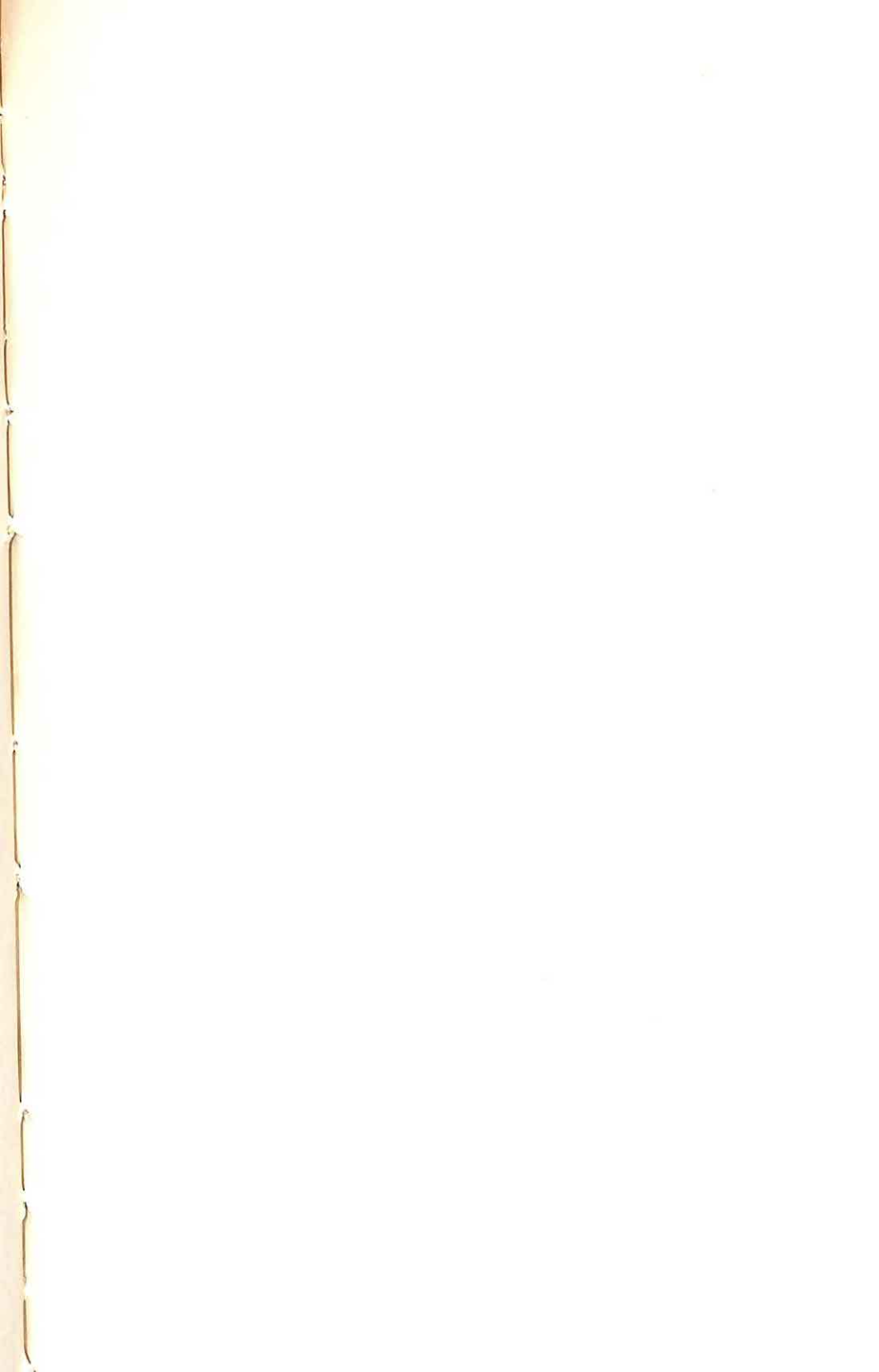
1. Some conceptual areas or hypotheses for adjectival pattern labels are more agreed upon by adults than by children;
2. Antonyms describing patterns often do not produce similar category usage, since one member of an antonym pair subsumes some distinct concepts, while the negation or contradiction of such concepts is not necessarily itself productive of distinct concepts of equivalent singularity;
3. The postulation of antonyms as being of equal and opposite semantic force is a convenient but probably artificial assumption in the description of simple patterns.

REFERENCES

1. ATTNEAVE, F. *Applications of Information Theory to Psychology*. New York: Holt-Dryden, 1959.
2. BRUNER, J., GOODNOW, J. J., & AUSTIN, G. A. *A Study of Thinking*. New York: Wiley, 1956.
3. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The Measurement of Meaning*. Urbana: Univ. Illinois, 1957.
4. RIDDERHEIM, D., & LEVY, N. A developmental study of the concept of symmetry. Paper read at Eastern Psychological Association, Philadelphia, April, 1958.

Department of Psychology
Rutgers University
New Brunswick, New Jersey







NEED REDUCTION AND PRIMARY REINFORCEMENT: VARIED GOAL SITUATIONS AS REINFORCEMENT IN MAZE LEARNING OF RATS*

Department of Psychology, University of Utah

BEATA JENCKS AND PAUL B. PORTER¹

A. INTRODUCTION AND PROBLEM

Investigations into Hull's hypothesis of identity between need reduction and primary reinforcement (8, Ch. VI) have reasonably well established the conclusion that need reduction can reinforce a habit even with the receptor system by-passed (e.g., 5, 19) and that primary reinforcement may be effected by incentives which are not need reducing (e.g., 16, 18). Studies concerning exploratory behavior, novelty, and curiosity are also pertinent to Hull's identity hypothesis. Manipulation is reinforcing to monkeys (7) and so is visual access to a more spacious environment (3). Rats can be roused to greater effort by stimuli of greater novelty (1, 2, 4, 12, 15, 20) or visual complexity (6). Further, rats have acquired both maze and bar-press habits when rewarded therefore by an opportunity to explore a relatively novel and somewhat more spacious environment (11, 13, 14).

Although several of the authors just cited used the word *novelty* to describe reinforcing stimuli, not always was the use strictly appropriate. If we are to know that novelty qua novelty is a reinforcer, the *S* must not be preacquainted with the character of the incentive lest there be objective anticipation, and he must gain the incentive through behavior performed in a familiar setting, i.e., the novel object must not be present. Incentives should not be consistently attractive to the *S* (like, e.g., a sweet taste), or consistently repulsive (like, e.g., an electric shock), nor should the behavior to be reinforced be provoked directly, "reflexly," by the situation (like, e.g., exploration).

The question now posed is: Will a rat learn and continue to perform on a maze when rewarded by the purely novel, with incentives chosen haphazardly with their variety limited only by the animal's ability to discriminate and the experimenter's ingenuity?

* Received in the Editorial Office on November 9, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.
¹ This article is based upon a thesis submitted by Beata Jencks to the University of Utah in 1959.

B. METHOD, SUBJECTS, AND PROCEDURE

The maze was a six-unit elevated *T*, copied from Miles (10). It terminated at a small table, 12" x 20", whereon the incentives might be placed.

Three groups of 10 naïve adult albino (Holtzman) rats were used. Each rat was separately caged and supplied with food (Purina Lab Chow checkers) at all times. All were given two weeks of preliminary training during which they were familiarized with one section of the maze and with the goal table. Further, throughout the preliminary training and the experiment itself, all rats were fed in their home cages once a day a paste of ground Purina Dog Chow and water. Becoming used to this regimen, the rats reduced their eating of dry food and tended to wait for the paste which was given about one hour after the daily handling and training.

In maze training, *Ss* were given one trial each day. One control group always found Purina paste on the goal table; the second control group found the table empty and was picked up immediately. The third (experimental) group found a new incentive every day: (*a*) the table carpeted with cotton, (*b*) a descending ladder replacing the goal table, (*c*) a dish of Purina paste, (*d*) being picked up as soon as they reached to goal, (*e*) a blast of air from an electric fan blowing across the table, (*f*) a strange rat, (*g*) a perfume stick, (*h*) a dark enclosure instead of the open table, (*i*) shredded paper piled deep, (*j*) nails strewn on the table, (*k*) no table, empty space, (*l*) the empty table.

Rats which did not reach the goal during the first three days of the experiment were eliminated. This left seven in the food group, eight in the pick-up group, and eight in the variety group.

C. RESULTS

The scores analyzed were the total number of cul entrances on Trials 2 through 12. A simple analysis of variance (9, Ch. XV) gave an *F* between groups of 8.2, with 2 and 20 *df*. The *t* values associated with comparisons of the several means appear in Table 1. The variety-rewarded group performed as well as the food-rewarded group; deviant, however, was the pick-up

TABLE 1
COMPARISON OF TOTAL TIME AND ERRORS SUMMED FOR TRIALS 3 THROUGH 12

Groups compared	<i>t</i> between Time	mean scores Errors	<i>df</i>
Food vs. Variety	0.4	0.3	13
Food vs. Pickup	2.5	3.3	13
Variety vs. Pickup	6.0	6.3	14

group. Figure 1 shows the means of the errors made by the three groups. The rather large number of errors made by the variety group on Day 6 is, we think, a consequence of the noise of the electric fan, from which the rats withdrew on Day 5.

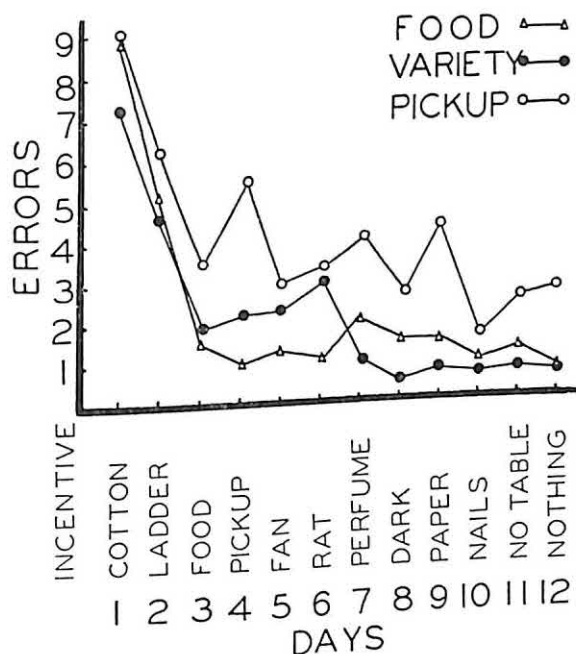


FIGURE 1
MEAN ERRORS PER TRIAL: INCENTIVES INDICATED ARE THOSE PRESENTED THE VARIETY GROUP

D. DISCUSSION

As intimated in the introduction, we intended to use a succession of rewards of little "intrinsic interest" to the rat and to preclude any objective anticipation. The rewards varied considerably in their individual attractiveness, but not many, we hazard, would be preferred to the food, once familiar. Nonetheless, as presented, the cumulative effect of the varied incentives was attractive—as attractive as food. If we appeal to an assumption that certain of the rewards were individually more attractive than the food and supported the behavior by way of partial reinforcement, it follows that the partial reinforcement resulting is unexpectedly effective.

To control anticipation is, of course, impossible; anticipation can only be made inappropriate. If we assume that there was objective anticipation, it

follows that there would have been parallel frustration. If present, such frustration should have hindered performance of the variety group.

The least demanding explanation of our results is Sheffield's (17) notion that reinforcement will be produced by any stimulus which elicits consummatory behavior—in this instance persisting investigation of the strange incentive. A generalization which seems to fit all the facts is that any stimulus whatsoever has a relative incentive value which is not zero, although it may sometimes be negative; only after the animal has had an opportunity to learn an appropriate response—to approach, to disregard, or to avoid, can we classify objects in the classic trilogy of positive, negative, and neutral. Stimuli are not neutral, and when they appear to be, it is because they have been habituated.

In the present experiment, in spite of relatively low deprivation, the reactivity of the food group and the variety group remained high. Variety seems to act very well as a reinforcer in rats, and the animals seem to be able to react to a "promise of novelty" at a spatial and temporal distance. Therefore, primary reinforcement should be so defined as to include novelty.

E. SUMMARY

Three groups of rats were given 12 trials on an elevated multiple *T*-maze: the first, 7 rats, was rewarded with a preferred food; the second, 8 rats, was just picked up at the goal; the third, 8 rats, found a different incentive each trial. The food and variety groups performed equally well, both reliably better than the pick-up group. These findings are interpreted as indicating that primary reinforcement must be so defined as to include novelty.

REFERENCES

1. BERLYNE, D. E. The arousal and satiation of perceptual curiosity in the rat. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 238-246.
2. BERLYNE, D. E., & SLATER, J. Perceptual curiosity, exploratory behavior, and maze learning. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 228-232.
3. BUTLER, R. A. Discrimination learning by rhesus monkeys to visual exploration motivation. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 95-98.
4. CHAPMAN, R. M., & LEVY, N. Hunger drive and reinforcing effect of novel stimuli. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 233-238.
5. COPPOCK, H. W., & CHAMBERS, R. M. Reinforcement of position preference by automatic intravenous injections of glucose. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 355-357.
6. DEMBER, W. N., EARL, R. W., & PARADISE, N. Response by rats to different stimulus complexity. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 514-518.
7. HARLOW, H. F., HARLOW, M. K., & MEYER, D. K. Learning motivated by a manipulation drive. *J. Exp. Psychol.*, 1950, **40**, 228-234.
8. HULL, C. L. *Principles of Behavior*. New York: Appleton-Century-Crofts, 1943.

9. MCNEMAR, Q. *Psychological Statistics* (2nd ed.). New York: Wiley, 1955.
10. MILES, W. R. The comparative learning of rats on elevated and alley mazes of the same pattern. *J. Comp. Psychol.*, 1930, **10**, 237-261.
11. MYERS, A. K., & MILLER, N. E. Failure to find a learned drive based on hunger; evidence for learning motivated by "exploration." *J. Comp. & Physiol. Psychol.*, 1954, **47**, 428-436.
12. MONTGOMERY, K. C. The effect of the hunger and thirst drives upon exploratory behavior. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 315-319.
13. ———. The rôle of the exploratory drive in learning. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 60-64.
14. ———. The relation between fear induced by novel stimulation and exploratory behavior. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 254-260.
15. MONTGOMERY, K. C., & MONKMAN, J. A. The relation between fear and exploratory behavior. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 132-136.
16. SHEFFIELD, F. D., & ROBY, T. B. Reward value of a non-nutritive sweet taste. *J. Comp. & Physiol. Psychol.*, 1950, **43**, 471-481.
17. SHEFFIELD, F. D., ROBY, T. B., & CAMPBELL, B. A. Drive reduction versus consummatory behavior as determinants of reinforcement. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 349-354.
18. SHEFFIELD, F. D., WULFF, J. J., & BACKER, R. Reward value of copulation without sex drive reduction. *J. Comp. & Physiol. Psychol.*, 1951, **44**, 3-8.
19. THOMSON, C. W., & PORTER, P. B. Need reduction and primary reinforcement: Maze learning by sodium-deprived rats for a subthreshold saline reward. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 281-287.
20. WILLIAMS, C. D., & KUCHTA, J. C. Exploratory behavior in two mazes with dissimilar alternatives. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 509-513.

Department of Psychology
University of Utah
Salt Lake City 12, Utah

EFFECT OF PSYCHOSTIMULI ON TECHNICAL PERSONNEL*

Hercules Powder Company, Inc.

HARRY TIPPER

A. SYNOPSIS

The following paper describes an experiment, in a real business situation but under relatively controlled conditions, to influence the performance of engineering and technical employees. An objective index of results was available. Personnel performance, measured by results, improved significantly. Methods used are applicable by any management concerned with specialist employees.

B. THE COMPANY

The business situation in which the experiment took place is representative of many fabricating and construction firms which render a technical service to industrial customers. In the present instance, the service was performed to the customers' designs and specifications. Actual performance of the work was done by mechanics under the supervision of engineers. Interpretation of customers' requirements was made by engineers or technicians, both for bidding purposes and for job performance. As is common in situations where competitive services cannot be differentiated in finite terms, orders were obtained almost invariably on a lowest-bid basis.

Ownership and management of the firm was in the hands of three graduate engineers and a business administration graduate who supervised accounting and office management functions. The first three acted respectively as general manager, production manager, and sales manager. The same executive personnel had been with the firm for more than 10 years, essentially in the same positions indicated by their titles.

The firm was located in the suburbs and, both by location and the natural tendencies of its managers, was relatively isolated from interchange of ideas with managers of competitive or comparable companies. Management procedures were highly developed and formalized, tending to inhibit contacts between managers and customers, or managers and employees in general. The entire office group with which we are concerned numbered only about 20.

* Received in the Editorial Office on November 19, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

C. THE SUBJECT GROUP

The part of this company given special attention is the engineering staff who analyzed customers' drawings and specifications, and prepared detailed estimates. This group comprised three to six men, mostly graduate engineers with five to 20 years experience.

Their analytical procedures were extremely detailed. They required the pricing of every piece of material to be purchased and estimates of the cost content of every operation to be performed by mechanics. As a guide to this estimating process, tables of unit costs had been made up from cost accounting records. Experience records on all previous jobs were kept available and consulted to provide a cross-check on the validity of each estimate being prepared. The finished estimate was compiled on a form which provided an automatic check-list of estimating steps so that checking of the estimate could be made by verifying individual items, as well as by comparison with over-all costs of previous jobs.

Nevertheless, individual judgement on the part of the estimator was still an important factor in the final price. It effected his concept of how the job should be done. It influenced the weight he gave each separate cost item. It was the principal element in determining the amount of contingency to be allowed for in preparing his final bid.

Hence, bids on the same job from competing companies frequently varied by more than 100 per cent where, on other occasions they might be almost identical. The very detailed estimating procedure had been developed in an effort to minimize this factor of judgement. Although this method, therefore, did not necessarily result in the correct bid, it had the effect of giving estimators and management confidence in their estimates: even if it were double the lowest, it was the "right" bid.

This procedure, of course, was extremely time-consuming. A single estimate would often require several man-months of analysis and compilation. Hence, the daily work tasks of this technical group comprised concentrated attention to a vigorous routine. Since the philosophy of the management was that the greater the productivity of estimates the more sales, management emphasis was laid on maintenance of this routine unrelentingly, with as little variation in activity as possible.

D. EFFECT OF THE EXPERIMENT

The experiment performed was directed at this latter point: that is, to provide variation in the level of attention and psychological focus. It was carried out over a period of seven months. Its effect can be measured against

a period of three years preceding it and 11 months following. The experiment can be isolated and treated as such because there was no other change in work procedure. Likewise, management policies remained stable over the entire period.

The effect of the experiment can be measured in terms of sales since, to make a sale, it was necessary to submit the low bid, usually against two or more competitors.

The gross effect of the experiment can be seen in the table of average monthly sales (Table 1).

TABLE 1

	143.
1953	179.
1954	143.
1955	98.
1956	266.
1957	104.
1958 (9 months)	

The test extended over only the first seven months of 1957. Comparing this term only of the same years, average monthly sales are indicated in Table 2.

TABLE 2

	131.
1953	285.
1954	175.
1955	83.
1956	409.
1957	106.
1958	

Thus, the test period produced volume more than double the average of these comparable periods.

The same results can be noted if we compare sales as a factor of estimating time (Table 3).

TABLE 3

	Est. man-mos/yr	Sales per est m/m
		45.9
1954	47	39.0
1955	44	33.8
1956	35	71.7
1957	44.5	20.5
1958 (9 mos)	45.5	

Examining the immediate time of the experiment, we find that sales per estimating man-month were as indicated in Table 4.

TABLE 4

1956	33.8
1957 (1st 7 mos.)	111.6
8/57 to 7/58	19.4

In the situation under consideration, we were breaking up a process which was designed to produce the maximum quantity of estimates and substituting for it one which would produce maximum results. We were reverting, in other words, from maximum mechanical efficiency in order to gain human factors. This drop-off in efficiency can be noted, also, in the figures.

The gross volume estimated per man-month went as indicated in Table 5.

TABLE 5

1953	
1954	700.
1955	851.
1956	862.
1957	876.
	535.

Comparing 1956 and 1957, we find that the total number of bids fell from 367 to 256. At the same time, gross volume bid in these years declined from 30,689 to 24,026, and number of successful bids from 52 to 37. However, the percentage of successful bids numerically held at 14 per cent both years, while the ratio of successes to dollar volume bid increased from 4 per cent to 12 per cent.

E. PSYCHOLOGICAL METHOD

The method used in obtaining these results was simply to inject into the work situation encouragement and incentive in normal human terms. There was no change in physical environment. The only changes in financial inducement were limited pay raises (in line with customary policy) and a moderate bonus on December 31, 1957. It is interesting that neither of these changes exerted any discernible beneficial influence on the sales results.

As in so many technical organizations, there was virtually no social contact between managers and staff during office hours. Even in coffee breaks and lunch periods, such social contact was negligible. In the same way, there was no attempt to include these engineers into management incentives for group performance. This technical group, therefore, was performing intensive mental effort in a psychological environment which was akin to that of a man on a treadmill.

In the particular situation, little could be changed and nothing could be effected permanently. During the test period, however, it was possible to make contact with the group, once or several times a day, to break into routine, to provide psychic stimuli to performance, and to exchange mutual encouragement over success. It is interesting to note that such small effort could have such important results. Certainly, they should encourage any management to give due regard to psychostimuli in their employee relations.

Hercules Powder Company, Inc.
Rocky Hill, New Jersey

PRODUCTION OF CROSS-TOLERANCE TO PSYCHOSIS- PRODUCING DOSES OF LYSERGIC ACID DIETHYLAMIDE AND PSILOCYBIN*

*Biological Laboratory, Cold Spring Harbor, New York; and South Oaks
Psychiatric Hospital, Amityville, New York*

H. A. ABRAMSON, A. ROLO, B. SKLAROFKY, AND J. STACHE¹

It has been previously demonstrated (4, 5) that 1-methyl-d-lysergic acid diethylamide (*MLD-41*) produces in man a marked cross-tolerance to d-lysergic acid diethylamide (*LSD-25*). *MLD-41* is less toxic than *LSD-25*, having about one-third of its psychotomimetic activity (1). Production of cross-tolerance to *LSD-25* by *MLD-41* occurs not only in nonpsychotic subjects but also in hospitalized schizophrenic patients (9). These data focus on the possibility that, if the schizophrenias are produced by a disturbance in biochemical mechanisms analogous to that resulting from the administration of mescaline, *LSD* and similar substances, comparatively non-toxic molecules might be administered therapeutically to produce tolerance to the endogenous chemicals that might originate schizophrenic states. In accordance with this concept, developed mathematically with Gorin (3), our present studies on cross-tolerance have been conducted not only using lysergic acid derivatives, but also employing a new psychotomimetic compound recently synthesized by Hofmann (8). He first isolated it from the Mexican mushroom, *Psilocybe*. The active principle, called *Psilocybin* by Hofmann, is a phosphoric acid ester of 4-hydroxy-dimethyltryptamine. The indole ring is possessed by *Psilocybin* in common with lysergic acid.

Exposure of Siamese fighting fish to concentrations of 100 μ g of *Psilocybin* in the outside liquid showed nothing worthy of note. However, when the fish are injected intraperitoneally with 25 μ g of *Psilocybin*, their reactions are identical with those produced by *LSD* itself. The nose up-tail down position, excitation and kink in the tail are all readily observed (2, 7). Thus, *LSD* and *Psilocybin* in the fish may affect the same enzyme systems as do the respiratory enzyme poisons (6, 10).

In man, studies on cross-tolerance were carried out as heretofore. The

* Received in the Editorial Office on November 20, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This investigation has been aided in part by a grant from the Josiah Macy, Jr. Foundation, New York, N. Y. We are indebted to Sandoz Pharmaceuticals for the supplies of *Psilocybin* and *LSD-25*.

effect of Psilocybin itself and production of cross-tolerance by *MLD*-41 and *LSD* on man were obtained by giving it to the same group of five nonpsychotic test subjects who have been used in the study of *LSD*-25 and its derivatives for the past five years. All the compounds were administered orally either in distilled water or, in the case of Psilocybin, in capsules containing up to 2 mg of Psilocybin per capsule. Development of cross-tolerance to *LSD* by Psilocybin or to Psilocybin by *MLD*-41 and *LSD* was achieved by administering these drugs at home for five to 12 days ahead of time in increasing doses. In general, the procedure followed was that described heretofore (4, 5). Although Psilocybin produces tolerance to itself, symptoms developed during the pretreatment period at home, and our subjects were cautioned not to drive or to engage in any outdoor occupation while under the influence of the drug. Our most sensitive subject to Psilocybin (C.G.) developed tolerance to 4 mg of Psilocybin within six days and was able to take 15 mg on one day, in divided doses, without serious symptoms. Another subject (M.Z.) took 103 mg of Psilocybin in increasing doses for 14 days and was able to take 6 mg of Psilocybin with only minor symptoms. The method of measurement of responses to all drugs was that described heretofore. A questionnaire consisting of a first part containing 47 questions and a second part used as a short mental rating test of nine reactions was employed. Positive responses to the questionnaire are added, irrespective of the intensity of the response. Because *MLD* can be more safely administered to subjects at home during the pretreatment period, we have used that in preference to *LSD*. *MLD* produces tolerance to itself very rapidly, and for this reason large doses of *MLD* may be safely administered. We have summarized for this preliminary report only those experiments on cross-tolerance (Table 1). Actually, in this particular series of experiments on the group of five (now six) subjects, 101 experiments were needed for obtaining controls, randomization, cross-over design, and placebo data. These data will be published in full elsewhere. Inspection of Table 1 shows that in all six subjects and in all 15 experiments that have been performed on production of cross-tolerance between *LSD*, *MLD*, and Psilocybin, cross-tolerance was observed, using our questionnaire method of observation in every case. It is our impression that Psilocybin is more effective in producing cross-tolerance to *LSD*, with the dosages and method employed, than the cross-tolerance produced by *MLD* and *LSD* to Psilocybin. Perhaps a better indication of our criteria of cross-tolerance production is given by the fractions following the words listed in Column 4 of Table 1. The numerator of each of these fractions is the total number of responses to the questionnaire when the test

TABLE 1

The effect of pretreatment of nonpsychotic subjects by 1-methyl lysergic acid diethylamide (*MLD-41*) and Psilocybin on test doses of Psilocybin and lysergic acid diethylamide (*LSD-25*) in a study of cross-tolerance production. Note that cross-tolerance was observed in all cases. The fractions in the fourth column are the ratio of questionnaire responses with and without pretreatment.

Subject	Pretreatment dose	Test dose	Tolerance to test dose
R.B.	1175 μ g MLD	6 mg PSI	Partial (2/5)
	4325 μ g MLD	8 mg PSI	Partial (2/9)
	103 mg PSI	75 μ g LSD	Almost complete (< 3/7)*
J.G.	4300 μ g MLD	8 mg PSI	Almost complete (1/29)
	67 mg PSI	75 μ g LSD	Almost complete (3/20)
	600 μ g LSD	8 mg PSI	Almost complete (2/29)
D.V.G.	1175 μ g MLD	6 mg PSI	Complete (0/10)
	4325 μ g MLD	8 mg PSI	Complete (0/14)
	103 mg PSI	75 μ g LSD	Complete (> 0/8)
M.Z.	815 μ g MLD	6 mg PSI	Almost complete (3/10)
	4325 μ g MLD	8 mg PSI	Partial (9/18)
	103 mg PSI	50 μ g LSD	Almost complete (3/27)
C.G.	50 mg PSI	50 μ g LSD	Almost complete (4/25)
P.B.	1175 μ g MLD	6 mg PSI	Complete (0/10)
	103 mg PSI	75 μ g LSD	Almost complete (< 8/18)

* Seven responses were to 50 μ g *LSD*.

dose was administered after pretreatment, while the denominator represents a value for the responses to the questionnaire given by control experiment with the test dose or less where no pretreatment dose had been administered the week before. Note that in every case, without exception, the value of the fraction is one-half or less. The data do not convey the remarkable blocking effect produced, for example, in Subject D.V.G., who has a most severe reaction to 75 μ g of *LSD-25* but who is able to take this dose of *LSD* with no symptoms whatsoever after pretreatment with Psilocybin. Table 1 illustrates that there is an important variation between subjects in the ability to develop cross-tolerance. What the mechanism is, we do not know, although Subjects J.G. and D.V.G. are the only two female members of the group.

It might be conjectured at first that the presence of the indole ring was primarily responsible for the production of cross-tolerance. This is not quantitatively borne out by our unpublished studies of cross-tolerance production by other similar compounds, *LAE*, *DAM*, and *UML*, which show a wide variation in effectiveness. Further experiments with d-lysergic acid, itself, and compounds similar to lysergic acid diethylamide and derivatives and congeners of the hydroxytryptamines are planned. In any event, these data indicate that there may be a common mechanism responsible for the

production of the induced psychotic states in man by compounds of this type. In other words, is the enzyme mechanism in the psychoses induced by *LSD* and Psilocybin the same? Further investigations of the nature of this mechanism, such as those we have begun, employing the respiratory enzyme poisons in Siamese fighting fish, may lead to a solution of the problem, not only of the genetic and enzyme processes connected with the experimental psychoses in man, but also those processes accountable for the schizophrenias themselves.

REFERENCES

1. ABRAMSON, H. A. Lysergic Acid Diethylamide (*LSD-25*): XXIX. The response index as a measure of threshold activity of psychotropic drugs in man. *J. of Psychol.*, 1959, **48**, 65-78.
2. ABRAMSON, H. A., & EVANS, L. T. Lysergic Acid Diethylamide (*LSD-25*): II. Psychobiological effects on the Siamese fighting fish. *Science*, 1954, **120**, 990-991.
3. ABRAMSON, H. A., JARVIK, M. E., GORIN, M. H., & HIRSCH, M. W. Lysergic Acid Diethylamide (*LSD-25*): XVII. Tolerance development and its relationship to a theory of psychosis. *J. of Psychol.*, 1956, **41**, 81-105.
4. ABRAMSON, H. A., SKLAROFKY, B., BARON, M. O., & FREMONT-SMITH, N. Production of tolerance to psychosis-producing doses of Lysergic Acid Diethylamide. *Science*, 1957, **125**, 1020.
5. ———. Lysergic Acid Diethylamide (*LSD-25*) antagonists: II. Development of tolerance in man to *LSD-25* by prior administration of *MLD-41* (1-Methyl-d-Lysergic Acid Diethylamide). *Arch. Neurol. & Psychiat.*, 1958, **79**, 201-207.
6. ABRAMSON, H. A., WEISS, B., & BARON, M. O. Comparison of effect of Lysergic Acid Diethylamide with potassium cyanide and other respiratory inhibitors on the Siamese fighting fish. *Nature*, 1958, **181**, 1136-1137.
7. EVANS, L. T., GERONIMUS, L. H., KORNETSKY, C., & ABRAMSON, H. A. Effect of ergot drugs on *Betta splendens*. *Science*, 1956, **123**, 26.
8. HOFMANN, A., FREY, A., OTT, H., PETRZILKA, T., & TROXLER, F. Konstitution-saufklärung und Synthese von Psilocybin. *Experientia*, 1958, **14**, 397.
9. LENNARD, H., & HEWITT, M. P. Unpublished data presented at conference on the use of *LSD* in psychotherapy. New York: Josiah Macy, Jr. Foundation, 1959.
10. WEISS, B., ABRAMSON, H. A., & BARON, M. O. Lysergic Acid Diethylamide (*LSD-25*): XXVII. Effect of potassium cyanide and other oxidase and respiratory inhibitors on the Siamese fighting fish. *Arch. Neurol. & Psychiat.*, 1958, **80**, 345-350.

Long Island Biological Association
Biological Laboratory
Cold Spring Harbor, New York

South Oaks Psychiatric Hospital
Amityville, New York

THE SPAN OF HUMAN DEVELOPMENT*

*Mental Hygiene Clinic of the Oakland VA Office and Department of Psychology of
Washington State University*

HERBERT C. ARCHIBALD AND CLARE W. THOMPSON^{1,2}

The nature of growth has engaged the attention of students in all branches of science, and the concept can as usefully be applied to societies or to crystals as to human beings. The common core is great enough so that the existence of a general theory of growth is suggested, and principles derived from lower levels of life can be used to understand higher levels. Continuous growth at a constant rate is rare in nature. Even in undifferentiated cells the volume of the mass increases more rapidly than that of the surface from which nutrition is obtained, producing a gradually declining rate of development toward a limit. All human growth turns eventually into declining rates. For example, the addition to the body of new cells shows deceleration. The development of mental ability follows the same general trend as physical development. It is also rapid at first, then slows down. The simplest growth curves exhibit an ogive shape with the absolute growth being small at first, rising to a maximum and then declining. Human growth, characterized by change in size and differentiation whether in anatomical structure or in physiological or psychological functions, is much more complicated. For example, the thymus regresses while other body systems are building up. Thus there are combinations of superimposed curves, a summation of many part processes not necessarily moving in the same direction.

Similar statements may be made about behavior. In early childhood we see simple and direct adjustive activity and a narrow range of manifestations. Indirect response patterns are evolved when direct ones are no longer sufficient. For example, academic achievement may compensate for physical inadequacy and bullying may obscure an underlying "inferiority complex." As development continues we see less direct adjustive activity and note that indirect activity becomes increasingly complex. This complexity of adult behavior is undoubtedly what has obscured the developmental picture in later years.

* Received in the Editorial Office on November 27, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ We are indebted to Read D. Tuddenham, consultant to Veterans Administration from the University of California, for his help in formulating some of the ideas presented here.

² Approved for publication by Veterans Administration Regional Office, San Francisco.

It is well established that development follows a specific sequence in childhood and infancy because the direct behavior of the child is easy to observe. The senior author interviewed several hundred mothers of children from one to 12 years of age. The number of "behavior" problems per child at any given age level was found to vary from one to six. For example, temper tantrums, fears, jealousy, and over-sensitiveness tend to increase up to age 4 or $4\frac{1}{2}$ and then to subside. Since these occur in over 50 per cent of normal preschool children, they cannot sensibly be regarded as neurotic behavior. Anorexia is most common from $2\frac{1}{2}$ to 5, soiling and enuresis decrease with age. Thumb sucking decreases as nail biting increases. Masturbation and restlessness in sleep show no trends with age in the preschool period. Some behaviors such as temper tantrums drop out, only to reappear later. Others rarely recur once they have dropped out. In childhood, the incidence of problems is greatest in the 5th and 12th years, with girls a little earlier. Negativism reaches its peak at $2\frac{1}{2}$. If parents can be persuaded to be patient this stage will not be greatly prolonged as the next is coöperativeness. It may be that the course of adult life follows just as regular a sequence, but to demonstrate this is much more difficult.

Our main thesis is that developmental psychology *should not* be seen as stopping at puberty. That, contrary to the commonly accepted stereotype in our culture,³ developmental factors continue to be of great importance throughout the entire life span. Biological time appears to be different from chronological time. We suggest that change is a constant factor. An individual progresses by fairly steady increments with acceleration and deceleration from time to time. The value judgments implied in using "development" to a certain point and then "decline" present a semantic fallacy.

One reason for the conventional view that development stops at puberty may be found in the history of developmental studies. Originally, curves of growth and decline were based upon cross-sectional samples. Next came longitudinal studies of young children, making available in the literature of the 1930's material on repeated measurements of a number of variables on children from birth through the preschool period. These grew into longitudinal curves of child development to which were appended cross-sectional curves of the adult population, typically showing decline from the age of puberty or early adulthood in the function portrayed. Only recently have the populations involved in longitudinal studies in the behavioral sciences reached early middle life. To the surprise of the investigators, these subjects did not show cessation of progress at puberty and subsequent de-

³ Example: Several middle-aged movie stars refuse to play the rôle of father.

cline, but rather are still increasing, at least in respect to cognitive functions (which can be described parametrically more easily than can emotional or relationship variables).

This history has two implications for our thesis. First, the point of view we are presenting as novel in this paper may be generally accepted by the time the subjects of longitudinal studies reach their fourth, fifth, and sixth decades.⁴ Second, we do not mean to imply that such things as traditional measures of mental age or physical growth will continue to increase throughout the entire life span. A man of 20 may be able to solve arithmetic problems faster than he will at 60, but at 60 he will be wiser. Who is to say that only the former may be called growth and the latter must be labelled decline. One study (1) has shown that test ability declines if the subtests contain visual and motor components and increases for subtests involving materials that depend on continued learning. Osler has asked: "What have bright eyes, red blood, quick breath, and taut muscles to do with philosophy?" The faculty of the Hastings Law School are all emeriti over age 65. Several jurists over 90 have successfully represented large corporations in recent years.

Of late, there is an increasing focus on the correlation of physiological and sociological findings. The point of view with which we are taking issue, in which development is seen as ending at puberty, is an excellent illustration of this. For what could be more physiological than development, nor more socially determined than viewing everything after early adulthood as decline rather than development. The Chinese would certainly find this incomprehensible, and could be expected to see all of life as preparing for the desirable condition of old age. Without belief in immortality life to many people seems short and brutish. In the absence of such a belief, youth and especially adolescence may well be overvalued. It is difficult to stand by and experience gradual loss of one's own teeth, vision, and muscular strength. To those who believe in post mortem reward, however, all of life may be seen as a stepping stone to happiness in another world and the "body beautiful" and youth become of relatively less importance.

Paradoxically, the value of a man used to be measured by his muscle power and that of his slaves. The discovery of the steam engine, representing the machine age, should have changed at least this emphasis. Only recently has muscle power given way. The use of coal, oil, gas, and uranium is seen as a stepping stone to a nuclear age. If loss of muscle power was the

⁴ In this connection, the Institute of Child Welfare at the University of California is changing its name to the Institute of Human Development.

reason for seeing aging as "decline," in the atomic age this would seem not to apply.

It should be obvious that there is a relationship between the study of the child and that of general laws of psychology which also apply to the adult. The trend to bring child psychology out of isolation in order to reintegrate it into general psychology is overdue. We think that the developmental (life span) point of view is applicable to all areas of psychology and suggest that this approach will provide a self consistent body of theory, cutting across traditional boundaries in psychology.

Developmental methodology has been applied to intelligence, perception, motor development, learning, social development, and genetic research. Although direct studies of children have contributed to theories of personality such efforts in the main are isolated and an experimental approach to the problems in this area has been minimal. Theories have in the most part been borrowed from clinical and animal psychology and by default "developmental psychology" is equated with child psychology.

There have been few longitudinal studies which have followed children with behavior problems in relation to adult psychiatric status. Some reports claim that most childhood problems disappear with maturation whereas others attribute improvement to treatment. A follow-up study (2) of patients referred to a child guidance clinic 30 years ago presents results on their present adjustment and compares them with a control group of adults who had no striking difficulties in school. The ex-patients were found to have a high rate of psychiatric illness as adults as compared with the normal controls. Of the ex-patients who were psychiatrically well as adults, most came from the group with neurotic problems as children. These children came from families of better socio-economic background than did the children with anti-social problems and delinquency. Class background, however, was not found to account for the greater proportion of psychiatrically ill adults among the latter subjects. The problems of serious delinquency for which children were seen in a psychiatric clinic in 1924 are now shown to have been the first symptoms of serious life-long psychiatric illness. The kinds of behavior problems these ex-patients exhibited as children were shown to be prognostic of their adult psychiatric status, although transient and reversible changes appeared alongside the enduring and irreversible changes that constitute developing.

Incidentally, the control group presented no cases of sociopathic personality or schizophrenia, although it was drawn largely from disadvantaged classes and a history of broken homes was found in one-third of the cases. Other

studies have shown that many well-adjusted adults have experienced markedly traumatic stresses in childhood. Psychiatrists and clinical psychologists regularly "find" such incidents in their cases and are all too prone to impute causal status to them. Relevant to many aspects covered in this paper is the tendency on the part of some observers to "cook the data," i.e., to see only that which coincides with their preconceptions. Thus, out of a great variety of factors one, such as breast feeding, may be selected as the sole causative factor of later character development.

The behavior patterns mentioned above include both adaptive and maladaptive characteristics, for development is not confined to characteristics considered admirable by our society. If the basic mechanism of biologic life is adjustive activity, the manner of adjustment depends upon the structure and past history of the organism. Adjustive mechanisms may be used for brief or prolonged periods. If recurrent, they are given names such as temper tantrums or negativism.

In the development of personality, maturity (development to optimal level) is seen as the overcoming of such "negative" factors as mistrust, shame, passivity, negativism, confusion of rôles, and isolation. The negative factors, however, are as much a part of the developmental process as are the positive ones which follow. Psychological growth is said to occur when such "positive" personality factors as trust, trustworthiness, persistence, consideration, relation to authority, imagination, ego identity, intimacy, and responsibility are integrated.

Throughout the course of differentiation and growth, basic interaction occurs between cells and their environment. Neither the developmental capacities of the cells nor the activity of external agents is sufficient to explain the final result. Both must always be taken into consideration. The organism through time changes with respect to susceptibility to modification. Apparently cells pass through a critical stage of growth during which they are most sensitive to environmental influences. Every cell contains the same genes and chromosomes. Presumably all have the same genetic potentiality early. For example, very early tissue transplants, such as thyroid to heart, continue the attributes of the site. Later transplants continue the attributes of the place of origin. Probably the same process is manifested on a macrocosmic scale.

There is considerable agreement that factors which affect development depend upon the phase in which they occur. Change in proportions of the body makes it possible for the child to function more efficiently as he grows older. These factors and phases are perhaps more important considerations in

prevention than in treatment. We must learn from observational data how to detect the disturbed child and apply psychotherapeutic steps appropriate to each age level and its group of disturbances.

Individual differences account for different reactions to similar stimuli. Whether we are dealing with heredity and developmental processes or a function of differing environmental pressures, they have meaning to the adjustive process. Establishing the relative importance of heredity and environment is not the most fruitful line of investigation into the problems of developmental psychology. To put the question in these terms clouds the issue. Differences exist in size and rates of growth, muscular development, nervous reactivity, sensory acuity, and so on. The individual who is at either end of the distribution curve has not only different built-in reactions to contend with, but he may also develop different attitudes and self concepts and may receive different treatment by the social environment. For example, the undersized boy will have a different attitude depending on whether his father was a star on the college football team or is a college professor.

Hereditary factors have been demonstrated in numerous diseases. One of the best known examples is hemophilia which occurs in males although it is transmitted through the female. This disease usually manifests itself in early childhood, as does hereditary ataxia, a rare developmental defect. Amaurotic idiocy appears to be a recessive Mendelian characteristic appearing chiefly in the Jewish group. In Schilder's disease, massive lesions appear in the subcortical white matter. Two or more children in the same family are often reported. The disease occurs most commonly in childhood but may develop at any age. In most of the above diseases of the central nervous system, it should be noted that the several members of the same family develop the disease at the same age. Sir William Gowers applied the name abiotrophy to such conditions. By this term he intended to imply that the nervous system was endowed with inadequate resistance or capacity to survive and that eventually certain structures cease to function and disintegrate.

It is not necessary for a pattern to appear in childhood in order to be hereditary, familial, or developmental. Huntington's chorea, a rare form of disease in the basal ganglia, occurs in middle aged and elderly persons, sometimes associated with mental disorder. The disease depends on a definite Mendelian genetic dominant determinate and constitutes an example of the interaction of heredity and development. There is no evidence of the influence of extra-organic factors.

The interaction of extra-organic factors and hereditary determinants is illustrated by combat neurosis. It has been found that neuropsychiatric dis-

orders occur in over half of the immediate families of these patients. Soldiers subjected to prolonged combat frequently present a syndrome of combat dreams, blackouts, headaches, acoustic hypersensitivity, and startle reactions. In these cases development appears to play little or no rôle.

The rôle of genetic factors in the production of mental disorder continues to be a source of controversy. There is general agreement that certain forms of mental disorder result from infections, trauma, toxins, congenital abnormalities, tumors, or vascular accidents. The pathology in such cases can be seen in the grey or white matter of the central nervous system. There is much less agreement when it comes to the vast majority of mental disorders, the so-called functional psychoses and psychoneuroses. Some experts, notably Kallman, view these disorders as primarily genetic in origin. Others tend to view them as almost entirely dependent upon environmental factors, usually stress sustained in early childhood, the earlier the more severe the reaction. More definite knowledge must await further research. In the meantime genetic and developmental factors should be considered along with physiological, psychological, social, and chemical factors.

Schizophrenia is rare under 15, suggesting that development is a factor, since a fairly narrow life span accounts for the onset of the large percentage of the cases. Mania and clinical depression are extremely rare in childhood. Some would explain this as dependent on development of the "superego." We would regard it as dependent on development of the *child*.

Nail biting, anorexia, overdependence, and excessive reserve occur more frequently with children with poor physiological status. As an example of the interrelationship between heredity and environmental factors, the undersized or oversized child frequently shows evidence of shyness or aggressiveness.

The similarity of content between developmental and learned behavior raises the question of how to decide which concept is applicable in a given instance. Conceptualizing case material in this framework depends upon whether something is innate as opposed to learned. This "rivalry" may not always be mutually exclusive. In many instances a symbiotic relationship may be seen between the initial developmental factors and interpersonal experiences which arouse them and which encourage continuation of efforts to gratify them. For example, parents who are progressively pleased as the child sits, creeps, stands, and walks contribute to his locomotive development, although this development must still be paced by his maturation. A six-year-old child is not taught calculus but he is expected to learn to read. The all too common phenomenon of the non-reader results when developmental factors (reading readiness) are disregarded. The slow maturing adolescent plays at the game of heterosexuality lacking a substantial developmental

substratum.⁵ We suggest that developmental and learned behavior are poles of what is basically a continuum.

It is doubtful if heterosexual adjustment is possible if all of the previous experiences of the individual are non-rewarding. Thus learning influences development. In the homosexual, development is normal chemically but the object is different, i.e., the gonads secrete testosterone but the person prefers men. In the child unacceptable sexual behavior is socially disapproved and makes him uncomfortable. This does not mean that the behavior drops out. Running into the street is also strongly disapproved but can be easily stopped by punishment compared to masturbation which is a modified outlet for something biologically impelled.

The developmental approach is not new in child psychology, pediatrics, child psychiatry, and psychoanalysis. Years ago Freud noted limitations of the direct observations of child psychologists. It is not easy to correlate reconstructive data from psychoanalysis with data of direct observation. In the case of sibling rivalry, for example, observation and questioning alone will not elicit the conflict. The child who says he just loves the new baby sister will treat roughly the doll baby in an anonymous play situation. Conversely, to understand a child's drawing or play, the observer may have to ask the child, who for example may have included in his picture both the house and the money with which to buy the house. Neither direct observation nor reconstruction alone give complete understanding.

Direct observation has led to the discarding of some earlier hypotheses and has proven useful in developing others. For example, Freud appears to have gotten his idea of repetition-compulsion from watching the game of an 18-month-old child who caused a spool to disappear and reappear whenever the mother left. On the other hand the reconstructive method enabled Freud to recognize developmental sequences in the child's life. The complexities of the problem can be further illustrated by mentioning the seduction hypothesis, which had to be abandoned with the consequent recognition of the importance of phantasy life. To balance this, one must say that conjecture as to the mental content of earliest infancy is open to question.

There are many unanswered questions. Should symptoms be treated directly? If symptoms are developmental, will they not drop out? How much can later experience mitigate earlier deviation? So far we lack even the statistical methods capable of analyzing developmental material, although with trend analysis and Cattell's "*t*" technique this situation is changing. Is

⁵ It is of interest that in recent years marriage frequently occurs before the individual establishes his psychological identity.

it helpful simply to stress that there exists a large developmental factor in all of human behavior? If so, how can this information best be transmitted to the professions, to teachers, and to parents?

Cause, content, meaning, description, and explanation are closely related in psychiatry and psychology. In this discussion of development it has undoubtedly been used as first one, then another of these constructs. Probably never is there a single cause, although one of the variables may be more important than the others. The developmental processes may be seen as a system of causes influencing learning. Learning in turn affects development, and the cycle is repeated once again. The purpose of this paper has been to bring to the attention of the reader the complex causal network of the developmental process and to emphasize that the developmental processes continue throughout the life cycle.

REFERENCES

1. CORSINI, R. J., & PASSETT, K. K. Intelligence and aging. *J. Genet. Psychol.*, 1953, **83**, 249-264.
2. O'NEAL, P., & ROBINS, L. N. The relation of childhood behavior problems to adult psychiatric status. *Amer. J. Psychiat.*, 1958, **114**, 961-969.

VA Office
1305 Franklin Street
Oakland 12, California

Department of Psychology
Washington State University
Pullman, Washington

\$20.00 per annum
\$15.00 per volume
\$7.50 per half volume

QUARTERLY
Two volumes per year
Immediate publication

April, 1960
Volume 49, Second Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

APRIL, 1960

(Manuscripts are printed in the order of final acceptance)

The auction value of certain risky situations	167
By PHILIP NOGEE AND BERNHARDT LIEBERMAN	
Weak-light reinforcement with and without control for response facilitation .	181
By W. F. CROWDER, W. P. WILKES, AND T. H. CROWDER	
A comparison of three rates of pulse onset and decline in producing critical flicker frequency	185
By S. HOWARD BARTLEY AND THOMAS M. NELSON	
Parent-child psychotherapy: a follow-up study comparing two techniques .	195
By E. LAKIN PHILLIPS	
Cutaneous communication: Absolute identification of electrical intensity level	203
By GLENN R. HAWKES	
The verbal power test of concept equivalents (VPT)	213
By E. FRANCESCO	
Screening solutions to upgrade quality: A new approach to problem solving under conditions of uncertainty	217
By NORMAN R. F. MAIER	

(OVER)

Copyright, 1960, by The Journal Press
Provincetown, Massachusetts, U. S. A.

Entered as second-class matter January 26, 1937, at the post-office at
Provincetown, Mass., under the Act of March 3, 1879

Modification of a commonly used source of electric shock	233
BY GUY H. MILES	
Epidemiology of mental health in college	235
BY RALPH M. RUST	
Time perceptions as affected by need tension	249
BY ROBERT D. MEADE	
✓ Test item difficulty and the reliability of item analysis methods	255
BY JAMES F. ADAMS	
Considerations involved in obtaining averages of <i>CFF</i> data	263
BY THOMAS M. NELSON, S. HOWARD BARTLEY, AND DORIS DEHARDT	
Maximum I_t for absolute identification of cutaneous electrical intensity level	279
BY GLENN R. HAWKES AND JOEL S. WARM	
Head-balance and sitting posture: An x-ray analysis	289
BY FRANK PIERCE JONES AND PHILIP F. M. GILLEY, JR.	
Problems in the quantitative evaluation of patients with psychoses of the senium	295
BY KENNETH LIFSHITZ	
The stimulus concept: <i>Animism</i> ?	305
BY LEO L. GLADIN	
The relationship of two measures of authoritarianism to the Minnesota Teacher Attitude Inventory	333
BY ALBERT J. KINGSTON AND GEORGE L. NEWSOME	
Behavioral research and intergroup communication	339
BY WILLIAM J. MACKINNON	
The student's need for an issue	349
BY MALCOLM H. ROBERTSON	

THE AUCTION VALUE OF CERTAIN RISKY SITUATIONS*

Department of Psychology, Boston University

PHILIP NOGEE AND BERNHARDT LIEBERMAN^{1,2}

A. INTRODUCTION

The studies reported in this article utilized the expected value model and present an extension of some recent work dealing with the concept of psychological probability. In particular, the studies yielded measures of psychological probability under varying conditions. In general form the model may be expressed as

$$EV = pW + (1 - p) L$$

where EV is the expected value, p is the probability of winning, W is the amount that may be won, $(1 - p)$ is the probability of loss, and L is the amount that may be lost, a negative quantity.

To obtain measures of psychological probability, Preston and Baratta (8) used a particular case of the model in which only the win quantity had a value; the loss was zero, when the Ss were considering the risky situation. The model then reduced to

$$EV = pW.$$

In their experiments, the Ss had to bid at auction for a gamble with stated probabilities and amounts that could be won. The Ss' bids were then taken as their estimates of the EV of the gamble they considered. Although the Ss' behavior did, in general, conform to the EV model, deviations did occur. The deviations were explained by hypothesizing a transformation of the scale of objective probabilities and the measures obtained were termed psychological probabilities. For low probability values, 0.01 and 0.05, individuals overestimated the likelihood of an event and for objective probabilities ranging from 0.25 to 0.99, underestimation occurred.

* Received in the Editorial Office on November 30, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This article is based on a portion of Lieberman's doctoral dissertation submitted to the Boston University Graduate School. We are happy to acknowledge the assistance of Professors Nathan Maccoby and Garry Margolius and the help of Drs. James F. Lawrence, Joseph Sacks, Donald Ramsdell, and Harold Goodglass of the Veterans Administration. The research was done at the Brockton and Boston Veterans Administration Hospitals and we are grateful to the research committee of the Brockton Hospital for its support of this work.

² Now a Postdoctoral Research Fellow of the National Institute of Mental Health in the Department of Social Relations, Harvard University.

Similar results were found in other studies. Griffith (5) found that horse race bettors behaved in a manner similar to Preston and Baratta's Ss. Sprowls (10) analyzed French, Mexican, and Spanish lotteries and found results that tended to confirm the findings that individuals overestimate the probability of winning when that objective probability is less than 0.20. Attneave (1) examined Ss' estimations of the occurrence of words in English text and his results indicated that individuals tended to overestimate the presence of those words with a low probability of occurrence and underestimate the presence of those words with a higher probability of occurrence.

Using a paired comparison method, Edwards (2) found that when Ss were faced with a choice among bets of equal positive expected value they preferred bets with a $4/8$ probability of winning and had an aversion to bets with a $6/8$ probability of loss. They preferred the alternative which had the lower probability of loss although those bets had larger amounts of loss. In a more recent study, Edwards (4) found that subjective probabilities of winning and losing are somewhat different. Ss tended to overestimate their chances of winning money while they had a fairly objective estimate of the chances of losing money.

Given this background of theoretical and empirical work the studies to be reported were undertaken to obtain measures of psychological probability under various conditions. The first study attempted to replicate the findings of Preston and Baratta using a different population, psychiatric aides rather than the original group of college students and faculty. In the second study similar procedures were used to obtain measures of psychological probability corresponding to $(1-p)$, or q values. A third study was done to obtain measures of psychological probability when amounts of play money were being lost.

B. METHODS AND PROCEDURE

1. Subjects

Forty men, psychiatric aides, employed at the Brockton Veterans Administration Hospital served as Ss in the studies. In the first study, 20 aides participated and 16 of these also took part in the second study. Twenty different aides were the Ss of the third study.

The Army General Classification Test (*AGCT*), Form *AM*, and the Woody McCall Mixed Fundamentals Arithmetic Test were administered to each S. The *AGCT* supplied some measure of the Ss' general intellectual abilities and the Woody McCall Arithmetic Test was used to eliminate the Ss who could not multiply correctly a decimal by an amount, the computations necessary to compute an *EP*.

The 20 aides who served as *Ss* in the first and second studies had the following characteristics: mean age 24.7 years, mean years of schooling 10.7, and an *AGCT* mean score of 114.6. The 20 aides who participated in the third study had a mean age of 26.7, mean educational level 11.7 years, and a mean *AGCT* score of 112.2.

2. *Materials Used*

The materials used in the studies consisted of: (a) a pair of dice, (b) \$86,000 in play money, (c) 49 stimulus cards that contained the bets, (d) sheets that explained the betting procedures, (e) cigarettes and candies, the prizes.

3. *Procedures*

The experimental procedures consisted of three experimental games. The first was a repetition of the essential elements of the Preston and Baratta procedures from which the estimates of psychological probability were derived. A bet with a given probability of winning and an amount of reward was auctioned off to two players. It was assumed that the high bid of the *Ss* represented the *EV* of that bet to them and could be used in the *EV* equation to obtain an estimate of psychological probability.

The first game was played with 49 cards: seven were used in the practice trials and 42 in the experimental trials. Each card offered the *Ss* the opportunity to win a number of dollars with a given probability of winning. Six amounts of play money were involved: 5, 50, 100, 250, 500, and 1,000 dollars. The seven probabilities used were 0.01, 0.05, 0.25, 0.50, 0.75, 0.95, and 0.99. The six amounts of money and seven probabilities yielded 42 combinations of probability and reward. An amount of money together with a probability value constituted a bet and each bet appeared once in the 42 cards. Every game was played by two players.

Each player was given \$4,000, his endowment or bankroll. The 42 cards containing the bets were placed in random order which was the same for all games. One set of cards was placed between the *Ss* together with the information sheets that explained the meaning of probabilities, listed the amounts of rewards, explained the bidding procedure, and listed the points corresponding to each probability value. All this material was in front of the *Ss* throughout the game.

The first practice card was turned up, read by *E*, and auctioned to the higher bidder; bidding started at one dollar. During the seven practice trials the *Ss* learned the details of bidding and the pay-offs and had the oppor-

tunity to ask questions about the meaning of probabilities and other details of the game.

After the bidding was completed for each bet, the higher bidder rolled a pair of dice. His task was to make a point, whose probability of occurrence was the same as the p value of the bet. If both S s passed, both refused to offer anything for the bet, it was not considered. If the point was made the higher bidder received the amount of the reward minus his bid, but if the point was not made, the higher bidder paid the amount he bid. This procedure was carried out for the seven practice and 42 experimental trials. For the practice trials the amounts won or lost were not kept by the E or S s. The money did "change hands" as an aid in learning the game.

At the end of the game the player with more money, the winner, received his choice of a pack of cigarettes or five candy bars. The loser was given a single candy bar for participating in the game.

The data recorded consisted of the highest bid of each S , which S took the bid, and whether he won or lost.

The primary object of the second study was to obtain measures of psychological values, Q_w corresponding to mathematical values q . In the second experimental game the probabilities given to the S s indicated the E 's likelihood of winning and E rolled the dice. The S 's likelihood of winning was the complement of E 's likelihood of winning. The S s had to compute this, without pencil and paper and use this value to estimate the EV of the bet.

As an example, in the first study, when S was faced with the following bet

0.25 100,

this meant that he had the opportunity to bid for a bet in which he had an 0.25 probability of winning \$100. The EV was the product of these values, \$25.00.

In the second study the same stimulus card meant that the S had the chance of bidding for a bet in which the E had an 0.25 likelihood of winning. The S 's chance was now the complement of 0.25 or 0.75, and the EV of the bet was \$75.00.

To obtain measures of Q_w , the procedures of the first game were changed in the following ways. The probabilities given were now said to be the E 's likelihood of winning and the E rolled the dice. If he made the winning point, the higher bidder paid the amount of his bid to the E . If the E did not make the point, he paid the higher bidder the amount of the prize minus the higher bid. Aside from these changes the procedures of the two games were the same.

The experimental procedures of the third study were designed to obtain

measures of psychological probability in a loss-type situation. Instead of bidding for a gamble that enabled the *Ss* to win, the *Ss* paid a forfeit to avoid the risk of a given loss. The *S* who lost the smaller amount of money and had more than his opponent at the end of the game was the winner and received the winner's prize.

In the third study a stimulus card indicated

$$0.25 \quad 100.$$

In this case the lower bidder rolled the dice trying not to obtain a point with a probability of occurrence of 0.25. If he did not make the point, he paid nothing. The higher bidder paid the amount of his high bid—as a forfeit to avoid the risk of losing the \$100. If both *Ss* passed, both had to roll the dice and take the risk. The amount of the highest bid was taken to be the *Ss*' estimate of the expected loss of the gamble. The above gamble had an *EV* of —\$75. The measures of psychological probability obtained in this third study, where the *Ss* paid to avoid the risk of loss, could be compared with similar measures obtained in the first study, where the *Ss* paid to obtain the chance of a gain.

The *Ss* were interviewed after they played the game. An attempt was made to have them verbalize how they played the game and to state what factors determined their bidding.

4. *Measures of Psychological Probability*

The measures of psychological probability obtained in the three studies were derived from the *Ss*' bids. In the first study, when the *Ss* were deciding what amount to offer, no loss quantity was involved and the model reduced to

$$EV = pW.$$

The estimates of psychological probability were obtained from the successful bids of the 10 pairs of *Ss* by saying

$$MB = P_w W \text{ and}$$

$$P_w = MB/W$$

where *MB* is the mean successful bid of 10 *Ss*, P_w is psychological probability in the first study (the win situation), and *W* is the amount that could be won. The mean bid is taken to be the *Ss*' estimate of the expected value and the concept of psychological probability P_w is used to rationalize the *Ss*' bidding behavior.

Similar analyses of the data of the second and third studies yielded measures of psychological probability, Q_w and P_l .

5. Differences from Preston and Baratta

There were some slight differences between the original study of Preston and Baratta and the present replication. In these studies the *Ss* threw two dice rather than three; we used information sheets and cards containing the bets to explain the experimental procedures rather than presenting the bets in story settings. In addition we used practice trials that were not present in the original study. These differences are discussed in greater detail elsewhere (6).

TABLE 1
MATHEMATICAL EXPECTATION (*EV*), MEAN SUCCESSFUL BIDS (*MB*), AND PSYCHOLOGICAL PROBABILITY (*P_w*), IN THE FIRST STUDY

<i>P</i>	Prize values						<i>P_w</i>	<i>P & B*</i>
	5	50	100	250	500	1000		
0.01 — <i>EV</i>	.05	.50	1.00	2.50	5.00	10.00		
<i>MB</i>	.70	1.00	1.40	4.70	2.10	4.80		
<i>P_w</i>	.140	.020	.014	.019	.004	.005	0.034	0.065
0.05 — <i>EV</i>	.25	2.50	5.00	12.50	25.00	50.00		
<i>MB</i>	.70	3.90	3.50	12.00	24.50	20.00		
<i>P_w</i>	.140	.078	.035	.048	.049	.020	0.062	0.092
0.25 — <i>EV</i>	1.25	12.50	25.00	62.50	125.00	250.00		
<i>MB</i>	1.20	12.70	14.00	43.90	58.60	117.00		
<i>P_w</i>	.240	.250	.140	.176	.117	.117	0.174	0.195
0.50 — <i>EV</i>	2.50	25.00	50.00	125.00	250.00	500.00		
<i>MB</i>	1.70	12.60	35.00	100.00	210.60	361.50		
<i>P_w</i>	.340	.252	.350	.400	.421	.362	0.354	0.443
0.75 — <i>EV</i>	3.75	37.50	75.00	187.50	375.00	750.00		
<i>MB</i>	2.60	36.60	72.90	201.00	417.50	697.00		
<i>P_w</i>	.520	.732	.729	.804	.835	.697	0.720	0.677
0.95 — <i>EV</i>	4.75	47.50	95.00	237.50	475.00	950.00		
<i>MB</i>	3.80	41.80	85.70	194.50	454.70	893.30		
<i>P_w</i>	.760	.836	.857	.778	.909	.893	0.839	0.734
0.99 — <i>EV</i>	4.95	49.50	99.00	247.50	495.00	990.00		
<i>MB</i>	3.80	43.00	87.50	227.70	473.00	746.50		
<i>P_w</i>	.760	.860	.875	.911	.946	.747	0.850	0.833

* Data obtained by Preston and Baratta.

C. RESULTS³

1. Study One

Table 1 contains the expected value, mean bid, and measures of psychological probability obtained from the 10 pairs of *Ss* in the first study. The measures of psychological probability obtained by Preston and Baratta are

³ The three studies reported here were replicated using groups of schizophrenic subjects. Results are reported in Lieberman (6).

included for comparison. From these data it is quite obvious that the general findings of Preston and Baratta were replicated. In the auction situation of the first study the *Ss* overestimated the likelihood of an event occurring when its objective probability of occurrence was 0.01 and 0.05 and underestimated its probability of occurring when probabilities ranged from 0.25 to 0.99.

Examination of the 42 measures of psychological probability in Table 1 indicated that overestimation occurred in four of the values that made up the psychological probability value corresponding to objective probability, 0.01; underestimation occurred in the bets where 0.01 was combined with the prize values of 500 and 1,000 dollars. Underestimation occurred in four of the six measures of psychological probability corresponding to objective probability, $p = 0.05$.

2. Study Two

The results of Study 2 presented in Table 2 indicate that the overestimation-underestimation phenomenon was again present. Psychological probability exceeded objective probability at $q = 0.01$ and 0.05 and was less than objective probability for q ranging from 0.25 to 0.99.

TABLE 2
MATHEMATICAL EXPECTATION (*EV*), MEAN SUCCESSFUL BIDS (*MB*), AND PSYCHOLOGICAL PROBABILITY (*Q_w*) IN THE SECOND STUDY

<i>p</i>		Prize values						<i>q</i>	<i>Q_w</i>
		5	50	100	250	500	1000		
0.01	— <i>EV</i>	4.95	49.50	99.00	247.50	495.00	990.00		
	— <i>MB</i>	4.00	44.25	97.63	241.13	494.88	971.13		
	— <i>Q_w</i>	.800	.885	.976	.965	.990	.971	.99	0.931
0.05	— <i>EV</i>	4.75	47.50	95.00	237.50	475.00	950.00		
	— <i>MB</i>	3.63	45.63	96.38	235.50	438.88	970.38		
	— <i>Q_w</i>	.726	.913	.964	.942	.878	.970	.95	0.899
0.25	— <i>EV</i>	3.75	37.50	75.00	187.50	375.00	750.00		
	— <i>MB</i>	2.50	33.50	74.50	160.13	387.75	737.50		
	— <i>Q_w</i>	.500	.670	.745	.641	.776	.738	.75	0.678
0.50	— <i>EV</i>	2.50	25.00	50.00	125.00	250.00	500.00		
	— <i>MB</i>	2.13	19.38	49.25	127.00	209.50	411.88		
	— <i>Q_w</i>	.426	.388	.493	.508	.419	.412	.50	0.441
0.75	— <i>EV</i>	1.25	12.50	25.00	62.50	125.00	250.00		
	— <i>MB</i>	1.00	10.63	13.25	68.13	82.38	63.75		
	— <i>Q_w</i>	.200	.213	.133	.273	.165	.064	.25	0.175
0.95	— <i>EV</i>	.25	2.50	5.00	12.50	25.00	50.00		
	— <i>MB</i>	.88	7.25	9.13	15.00	42.75	52.13		
	— <i>Q_w</i>	.176	.145	.091	.060	.086	.052	.05	0.102
0.99	— <i>EV</i>	.05	.50	1.00	2.50	5.00	10.00		
	— <i>MB</i>	.63	.88	7.88	8.13	9.38	11.38		
	— <i>Q_w</i>	.126	.018	.079	.033	.019	.011	.01	0.048

Figure 1 presents the results of the first and second studies graphically. Psychological probability values are similar in both studies; overestimation occurred at probability values 0.01 and 0.05, and underestimation occurred at objective probability values 0.25, 0.50, 0.75, 0.95, and 0.99.

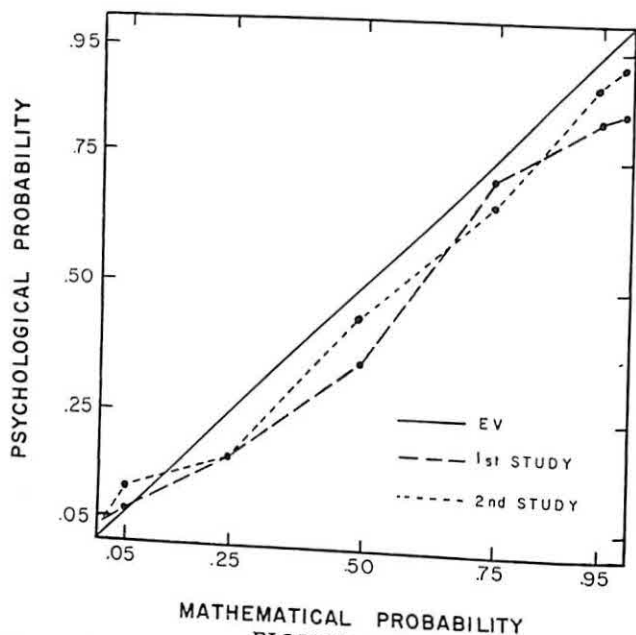


FIGURE 1
FUNCTIONAL RELATIONSHIP BETWEEN PSYCHOLOGICAL AND MATHEMATICAL PROBABILITY
IN THE FIRST AND SECOND STUDIES

3. Study Three

The results of the third study presented in Table 3 indicate that behavior in a loss type situation appears to differ from behavior in a win type situation. The overbidding effect at $p = 0.01$ and 0.05 was slight if it existed at all. Underbidding was characteristic at objective probability values 0.25 , 0.50 , 0.95 , and 0.99 and a slight overbidding effect may have been present at $p = 0.75$.

Figure 2 allows graphical comparison of the results of the first study and third study. At every objective probability value the Ss' bidding conformed more closely to the EV model in the loss type situation than the win type situation. Although some of the differences were slight, the results were consistent. Edwards' finding that Ss have a more objective estimate of their chances of losing amounts of money than their chances of winning such amounts appears to have been confirmed. In Edwards' study the Ss were in

TABLE 3
MATHEMATICAL EXPECTATION (EV), MEAN SUCCESSFUL BIDS (MB), AND PSYCHOLOGICAL PROBABILITY (P_l) IN THE THIRD STUDY

p		Prize values					P_l	
		5	50	100	250	500		1000
0.01	EV	.05	.50	1.00	2.50	5.00	10.00	0.018
	MB	.10	2.50	.60	1.50	1.10	25.10	
	P_l	.020	.050	.006	.006	.002	.025	
0.05	EV	.25	2.50	5.00	12.50	25.00	50.00	0.050
	MB	.40	8.10	1.30	6.20	8.70	3.80	
	P_l	.080	.162	.013	.025	.017	.004	
0.25	EV	1.25	12.50	25.00	62.50	125.00	250.00	0.197
	MB	1.60	8.90	12.10	36.70	83.00	248.60	
	P_l	.320	.178	.121	.147	.166	.249	
0.50	EV	2.50	25.00	50.00	125.00	250.00	500.00	0.411
	MB	1.50	22.60	44.60	119.60	249.90	290.60	
	P_l	.300	.452	.446	.478	.500	.291	
0.75	EV	3.75	37.50	75.00	187.50	375.00	750.00	0.767
	MB	2.90	39.50	78.50	215.30	405.50	773.10	
	P_l	.580	.790	.785	.861	.811	.773	
0.95	EV	4.75	47.50	95.00	237.50	475.00	950.00	0.932
	MB	4.60	41.90	93.70	245.70	482.40	948.80	
	P_l	.920	.838	.937	.983	.965	.949	
0.99	EV	4.95	49.50	99.00	247.50	495.00	990.00	0.930
	MB	3.30	48.90	97.00	248.50	495.20	989.70	
	P_l	.660	.978	.970	.994	.990	.990	

an individual choice situation while in these studies the S s were in an auction game situation (4). It appears that in both loss situations the S s' estimations of objective probabilities were more realistic than their estimates in win type situations.

4. Summation of Psychological Probabilities

By summing appropriate probability values of the first and second studies and corresponding psychological probability values obtained from the second and third studies, it was possible to obtain two sets of data dealing with the question of whether psychological probabilities sum to 1 as objective probabilities do.

Figure 3 summarizes the results obtained by summing the appropriate values of P and Q . The value of $p + q$ for all given combinations is 1. The various combinations of $P_w + Q_w$ yielded values that ranged from 0.965 to 0.745. The low point of the curve was at $p = q = 0.50$. The summation of $P_l + Q_w$ values yielded somewhat different results. The lowest point of the curve occurred where $p = q = 0.50$; the highest value occurred at $p = 0.95$,

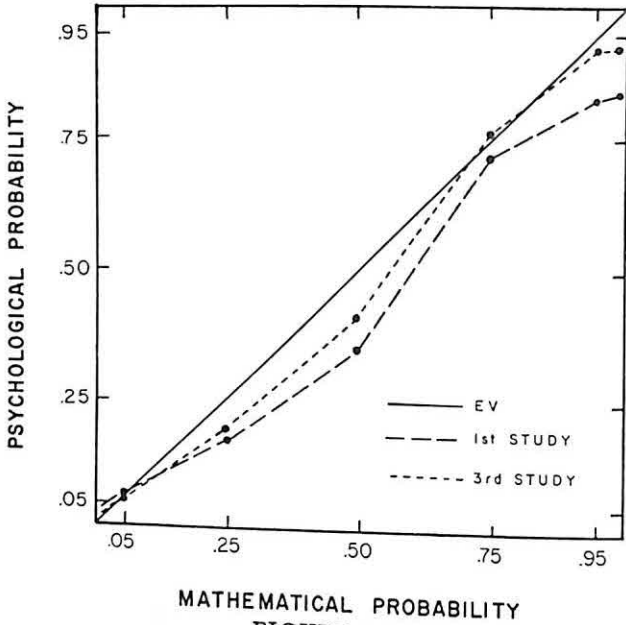


FIGURE 2
FUNCTIONAL RELATIONSHIP BETWEEN PSYCHOLOGICAL AND MATHEMATICAL PROBABILITY
IN THE FIRST AND THIRD STUDIES

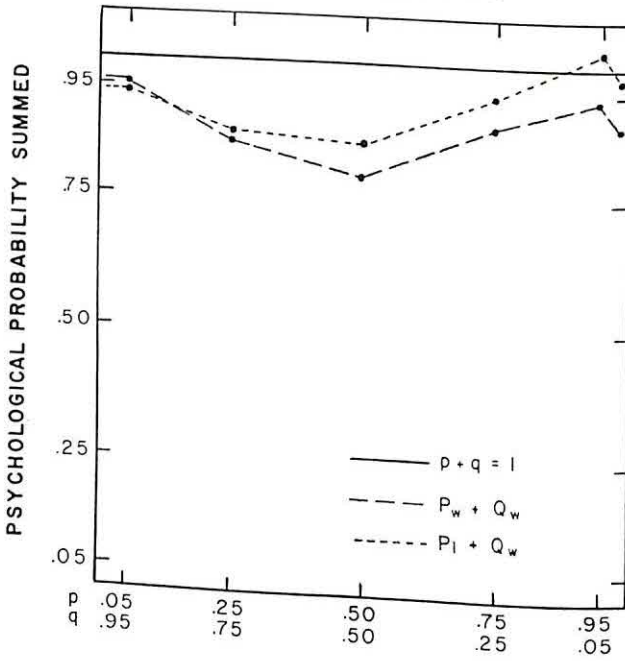


FIGURE 3
SUMMATION OF PSYCHOLOGICAL PROBABILITIES

$q = 0.05$. The elevation of the curve at this point resulted from the small underestimation of P_l and the overestimation of Q_w . The differences between the two curves may be understood as reflecting differences that occur when two estimates of psychological probability obtained from win situations are compared with estimates obtained from one win and one loss situation.

D. DISCUSSION

The results of the first study make it quite apparent that the general findings of Preston and Baratta were replicated in spite of certain slight changes in the experimental procedures and differences between the original group and the present group of Ss. The original group consisted of college students and faculty members; the present group consisted of psychiatric aides employed at a Veterans Administration hospital. The measures of psychological probability obtained were similar despite obvious differences between the two groups of Ss in educational level, intellectual abilities, and socio-economic class.

The second study gives some information about the abilities of individuals to make EV estimations. The Ss were given the E 's likelihood of winning, p . An analysis of the data suggests that with this information the Ss estimated q and used this estimate to obtain an expected value. Analyses of the results of the three studies indicated that Ss appeared able to estimate expected values under varying conditions and that they were able to combine probability and prize values to obtain estimates of appropriate bids. The bidding behavior of the Ss follows, fairly closely, the behavior called for by the objective model, although there were consistent deviations. However, the statements made by the Ss when they were interviewed at the conclusion of the game indicated that they could not give an accurate description of their behavior. Few stated that they combined probability and prize value to obtain their bids. Most gave personalistic or "superstitious" reasons for their bidding. Typical comments were, "I played hunches" or "I bid high when I felt I was going to make the point."

Preston and Baratta considered the possibility that the differing values of the prize were responsible for the variations from mathematical expectation. However, they rejected this notion pointing out that no matter what the prize, overbidding was associated with low probability values and underbidding was associated with higher probability values. A more detailed analysis of the data of these studies done elsewhere indicates that for the probability values ranging from 0.25 to 0.99 the size of the prize does not seem to affect the bidding of the Ss (6). For these probability values under-

estimation generally occurred no matter what the prize value. However, for probability values 0.01 and 0.05 the results are not so conclusive. Examination of Tables 1 and 3 indicates that in the first study there is some underbidding associated with the low probability values. Similar findings were obtained in the third study. However the results do not appear to be conclusive, and specific research would have to be done to determine conclusively whether the overbidding effect is present at the low probability values no matter what the prize value.

Edwards (3), and Mosteller and Nogee (7) discussed the summation of psychological (or subjective) probabilities. The meaning of the summation of objective probabilities is clear and derives from the conventional mathematical definition of probability. However, the significance of the summation of psychological probabilities is not at all unambiguous. It is possible to raise the question, what is the sum of psychological probabilities whose corresponding objective probabilities sum to 1? There are numerous ways of obtaining measures of psychological probability to answer this question. The present work adopted the general method suggested by Ramsey (9); a betting method was used to obtain the estimates.

The second study yielded measures of psychological probability corresponding to q , or $1-p$. By summing these values with appropriate values obtained from the first study, the results presented in Figure 3 were obtained. The results of this summation process yielded a curve whose points were all less than 1. This finding reflected the consistent underestimation of the probability values ranging from 0.25 to 0.99 present in the results of the first and second studies. The summation of the values of the second and third studies produced somewhat similar, but certainly not identical, results. Six of the seven values obtained were less than 1.

The empirical import of these results is not completely clear. However, these results illustrate the point that there is no reason to assume that psychological or subjective probabilities should sum to 1, although in certain situations they may. Since S s often do not behave precisely as the expected value model prescribes and since psychological probabilities often differ from objective probabilities, summation of psychological probabilities can readily yield sums other than 1. It may be that S s' psychological probabilities refer to specific events with which they are concerned. They may not make any precise estimate of alternative occurrences. This, of course, would be a marked departure from the objective model where all alternatives and associated probabilities are considered simultaneously.

E. SUMMARY

Three studies were conducted to obtain measures of psychological probability in win and loss type situations. The first study was a replication of the work of Preston and Baratta and yielded measures of psychological probability in a win type situation. The second study yielded measures of psychological ($1-p$), while the third gave measures of psychological probability in a loss type situation.

The results of the first study tend to confirm Preston and Baratta's original findings that Ss overestimate the likelihood of objective probabilities of 0.01 and 0.05 and underestimate probabilities ranging from 0.25 to 0.99.

The results of the second study indicate that Ss are capable of estimating q given p , a probability value, and then are able to use the q quantity in an expected value computation with considerable consistency. In addition, the overestimation-underestimation effect is present; Ss tend to overestimate q when that value is 0.01 and 0.05 and underestimate the q value when it ranges from 0.25 to 0.99.

The measures of psychological probability obtained from the third study, when compared with the results of the first study, indicated that Ss' estimations of objective probability in a loss situation are more realistic than their estimates in a win type situation. This result is consistent with previous findings.

REFERENCES

1. ATTNEAVE, F. Psychological probability as a function of experienced frequency. *J. Exper. Psychol.*, 1953, **46**, 81-86.
2. EDWARDS, W. Probability preferences in gambling. *Amer. J. Psychol.*, 1953, **66**, 349-364.
3. ———. The theory of decision making. *Psychol. Bull.*, 1954, **51**, 380-417.
4. ———. The prediction of decisions among bets. *J. Exper. Psychol.*, 1955, **50**, 201-214.
5. GRIFFITH, R. M. Odds adjustments by American horserace bettors. *Amer. J. Psychol.*, 1949, **62**, 290-294.
6. LIEBERMAN, B. The auction value of uncertain outcomes in win and loss type situations. (Unpublished Doctoral Dissertation, Boston University, 1958.)
7. MOSTELLER, F., & NOGEE, P. An experimental measurement of utility. *J. Polit. Econ.*, 1951, **59**, 371-404.
8. PRESTON, M., & BARATTA, P. An experimental study of the auction-value of an uncertain outcome. *Amer. J. Psychol.*, 1948, **61**, 183-193.
9. RAMSEY, F. Truth and probability. In F. Ramsey, *The Foundation of Mathematics and Other Logical Essays*. New York: Harcourt Brace, 1931.
10. SPROWLS, R. C. Psychological-mathematical probability in relationships of lottery gambles. *Amer. J. Psychol.*, 1953, **66**, 126-130.

Department of Social Relations
Harvard University
Emerson Hall
Cambridge 38, Massachusetts

WEAK-LIGHT REINFORCEMENT WITH AND WITHOUT CONTROL FOR RESPONSE FACILITATION*

Department of Psychology, University of Mississippi

W. F. CROWDER, W. P. WILKES, AND T. H. CROWDER¹

A. INTRODUCTION

Many recent experiments have reported that the onset of a weak light can serve to augment bar pressing, i.e., that the rate of pressing increases when each press brings the light. As has been pointed out (2), this increased responding need not imply reinforcement by the light. Instead, the light might merely facilitate the lever response, e.g., by making the *S* more active. The present study is concerned with the control of such facilitating effects. In each of the two experiments to be reported, two measures of the reinforcing power of light onset were obtained: (a) rate of pressing when pressing was followed by light; (b) rate of pressing when light was no longer given. The latter measure is believed to be free of facilitation effects and so may provide a "purer" index of reinforcement.

B. EXPERIMENT I

1. Method

a. Subjects and apparatus. The *Ss* were 11 naïve male Holtzman albino rats, approximately 85 days old. The apparatus was a two-lever Skinner box made from a wire-mesh living cage, 9 in. wide by 14 in. long by 7 in. high. It was shielded from sound and light by a large cover box. The two identical sheet-steel levers were 2 in. wide, 3½ in. apart, and 4½ in. above the floor. The *S*'s end of a lever was rolled into a 1-in. cylinder which extended 1½ in. into the box. The lever movement was ⅛ in., extremely quiet, and required a force of 11 gms. A 6-w., 110-v. clear glass lamp was mounted midway between and 5 in. above the levers. A resistor reduced the lamp voltage to 95 volts.

b. Procedure. The experiment lasted three days: one day each of operant-level testing, lever training, and testing for retention of lever training.

* Received in the Editorial Office on December 1, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The authors wish to thank Drs. W. A. Wilbanks and J. B. Wolfe for their critical reading of this paper. The study was supported in part by a grant from the Committee on Faculty Research.

(1). *Operant level.* The purpose of this phase was to measure the initial lever biases. Each *S* was left in the apparatus for 60 min. and the responses made on each lever were registered on magnetic counters. No light was presented. The measure of bias was the difference in number of responses made on the two levers during the last 50 minutes. The *S* with the greatest bias was discarded and the data reported are based on the remaining 10 *Ss*.

(2). *Lever training.* Again each *S* spent 60 min. in the Skinner box. Each response made on one of the levers turned on the 6-w. lamp for one second; pressing the other lever did not produce the light. Five randomly selected *Ss* were reinforced on their preferred levers; the other five were trained against their initial biases.

(3). *Retention test.* The procedure on this day was like that on the operant-level day, i.e., the *Ss*' responses were recorded for 60 min. without light.

2. Results

During lever training, the mean number of responses on the reinforced bar was 52.5; on the other bar, 38.1. This difference was significant at the .001 level according to the *t* test ($t = 5.81$). On the retention test day, the mean number of responses on the previously reinforced bar was 23.2; on the other bar, 23.5. This difference did not approach significance ($t = -0.1$). Thus the preference obtained during lever-training was entirely absent on the following day, when facilitation could not occur.

C. EXPERIMENT II

The procedure here differed from that of Experiment I in two ways: (a) instead of receiving a fixed period in the Skinner box, the *Ss* were removed after a total of 30 presses had been made on both bars, together; (b) the operant-level day was deleted and lever training took place on the first day of the experiment.

1. Method

a. *Subjects and apparatus.* The *Ss* were 20 naïve male Holtzman albino rats, about 70 days old. The apparatus was the same as in the preceding experiment.

b. *Procedure.* The experiment was completed in two days: one day of lever training and one day of testing for retention.

(1). *Lever training.* During the first 10 min., responses were neither recorded nor reinforced. Then, for 10 randomly selected *Ss*, the 1-sec. light was presented after each press of the left bar; for the other 10 *Ss*, after each

press of the right bar. The *S* was removed from the Skinner box as soon as possible after a total 30 responses had been made on the two bars.

(2). *Retention test.* The *S* was left in the Skinner box for 30 min. and responses to each lever were recorded but were not reinforced.

2. Results

During the lever training, the mean number of responses on the reinforced bar was 19.55; on the other bar, 13.80. The difference was significant at the .001 level according to the *t* test ($t = 4.53$). On the retention test day, the mean number of presses of the previously reinforced bar was 29.6; of the other bar, 24.3. This difference was significant at the .05 level ($t = 2.13$). The second experiment, unlike the first, thus found reinforcement effects even when response facilitation was eliminated.

D. DISCUSSION

The two indices of reinforcement are seen to give only partly similar results. According to the first measure (responses during lever training), both experiments found moderate reinforcement effects, with about 40 per cent more responses on the reinforced bar. According to the retention test, the first experiment found no evidence of reinforcement while the second experiment found a significant but rather small (22 per cent) preference for the reinforced lever. Only the retention test is thought to be entirely free of the effects of response facilitation by light.

Assuming, then, that retention is the more valid measure of learning, the present study does not provide much evidence that light is reinforcing. A possible explanation for this might be the slowness with which the two-bar task is learned when "weak" reinforcers are given. Forgays and Levin (1) used a two-bar Skinner box in a study of light reinforcement. They gave 10 days of lever training and found differential pressing in favor of the reinforced lever to reach its maximum on the ninth day. The present apparatus was also used in two secondary reinforcement experiments (3). In neither experiment was the bar preference retained for 24 hours.

E. SUMMARY

This study compared two measures of weak-light reinforcement in the Skinner box: performance during the time that the light was being used to condition lever-pressing; and performance on the day following lever training, with the light no longer given. The latter measure is held to be free of an important artifact, namely, any tendency that the light might have to evoke or to facilitate the lever response. Two experiments are reported in

which the light was used to produce differential pressing of two levers. On the training day, there occurred in both experiments about 40 per cent more presses of the reinforced lever. This could have reflected reinforcement, facilitation, or both. On the retention test, however, only one experiment found any evidence that the light was reinforcing. There, 22 per cent more responses were made on the previously reinforced bar. Since the two indices thus do not always give qualitatively similar results and since only the "retention" measure seems to be entirely free of facilitation effects, it would seem advisable to use the latter measure whenever the presence of the facilitation artifact is suspected.

The failure to find consistent reinforcement effects may have been due, it was suggested, to the difficulty of learning to press one of a pair of identical levers.

REFERENCES

1. FORGAYS, D. G., & LEVIN, H. Discrimination and reversal learning as a function of change of sensory stimulation. *J. Comp. & Physiol. Psychol.*, 1959, **52**, 191-194.
2. KLING, J. W., HOROWITZ, L., & DELHAGEN, J. E. Light as a positive reinforcer for rat responding. *Psychol. Rep.*, 1956, **2**, 337-340.
3. WILKES, W. P., & CROWDER, W. F. Secondary reinforcement with and without control for response facilitation. *J. of Psychol.*, 1960, **49**, 83-86.

Department of Psychology
University of Mississippi
University, Mississippi

A COMPARISON OF THREE RATES OF PULSE ONSET AND DECLINE IN PRODUCING CRITICAL FLICKER FREQUENCY*

Department of Psychology, Michigan State University

S. HOWARD BARTLEY AND THOMAS M. NELSON¹

A. PROBLEM

One of the stimulus factors considered crucial in determining response characters of the visual system to repeated stimulation is the rate of onset and decline of the photic pulses used. The most usual device for providing trains of photic pulses has been the episcotister, a revolving disk with one or more open sectors which interrupt a steady beam. Studies using this device are too numerous to list. A number of more recent studies have employed electronic circuits which activate "instantaneous" lamps and provide square waves of desired durations and frequencies of repetition (2, 3, 5, 6, 7, 8, 9, 10, 11).

While it may be thought that square-wave pulses embody the time-characteristics for receptor activation most easily interpretable, the usefulness of episcotister-produced pulses still exists. Much needs to be done to relate present day problems and findings to the accumulated literature on flicker which involved episcotisters. Also, the use of episcotisters manipulates the overall impingement conditions in certain waves not producible by present electronic pulsers. For example, the very movement of the episcotister blade across the target field, an undesirable event in some cases, provides a rich source of visual phenomena in others.

In the early days of the analysis of sensory response to photic stimulation, some attention was paid to disk design and other instrumental factors. Since that time, much less attention has been given to the effects of disk design and associated matters. Filehne (4), Schenck (10), and Schenck and Just (11) found that varying the number of sectors in episcotister disks influenced results, but did not present curves to show how this factor was related to others such as pulse intensity and *PCF* (pulse-to-cycle fraction).

Schenck and others of his time seemed to be set toward explaining the differences they found by use of the more common concerns of that day,

* Received in the Editorial Office on December 13, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The present study is a part of the work being done under Research Grant NSF-G-5821 from the National Science Foundation.

such as brightness contrast or induction, or to general notions of retinal lag and inertia. In using terms such as brightness contrast, they did not see that they were using sensory phenomena to explain sensory phenomena, rather than using understandings of body mechanisms for this purpose. Of course, they had much less knowledge of photochemistry and neurophysiology at their command than we have nowadays. What they had to say regarding retinal characteristics, such as inertia and lag, was not then specifiable in concrete processes and, hence, is of little contemporary significance.

One can use disks having various numbers of open and solid sectors. Obviously, such disks must be rotated at speeds in accord with the number of sectors to provide the same number of photic pulses per second. The chief difference the number of sectors in a disk makes is in the taper of the onset and decline of the pulses produced. The more sectors, the slower the disk needs to revolve to provide a given number of pulses per second, and, therefore, the more gradual the taper. With few open and solid sectors, the taper of the pulses produced is so brief that the relative durations of pulses and the cycles of which they are a part follow roughly the *PCF*'s indicated by the ratios of open to closed fractions of the disks themselves. But as numbers of sectors decrease, the tapers of onset and decline encroach both on the times during which the pulses are at full intensity and the times between pulses. At some point in the multiplication of sectors, the times during which the pulses exist at full intensity is zero, and the intervals between pulses are also zero. Still further multiplication of sectors would produce only oscillations of intensity with no fully unilluminated periods whatsoever.

Another factor in determining steepness of taper is the width of the slit in the lamp house which controls the width of the beam which the disk must interrupt. The broader the slit, the greater the fraction of the disk's rotation, and the greater the time taken to obstruct or expose the beam. Thus, the greater the taper of pulse onset and decline.

A third factor is the size of disk. The greater the diameter of the disk, the faster the open and closed sectors pass the beam for a given rate of revolution. These factors were obvious for the older workers mentioned, but the relation between pulse taper and the setting up of impulses in the receptors was not a thing of concern. To account for the ultimate sensory outcome, one would have to know just how various pulse tapers modify the rates of photochemical reaction in the receptors and, in turn, set up efferent impulses in the neuroretina. Probably enough is known about retinal photochemistry to suggest, in principle, the relations between rate of pulse onset

and decline and photochemical processes, but a satisfactory delineation of the receptor behavior under the conditions being discussed has not yet been achieved.

One would expect that an abrupt pulse would activate the available receptors most nearly simultaneously, and that various rates of taper would distribute the receptor bursts in time, at least at the beginning of the presentation of the pulse train. With repetition of pulses, the distributions would take on added complexity. Repeating pulse deliveries at high rates often involves such short pulses that the Bunsen-Roscoe law (the reciprocity between pulse duration and intensity to produce a constant end-result) is brought into operation, and this is still another complexity.

The procedure problem in the present study was simply (a) to use several disks, each with a different number of sectors, keeping the *PCF* constant, and varying pulse intensity to determine *CFF*; and (b) to vary *PCF* to see what rôle it played in determining *CFF* as number of sectors was varied.

B. PROCEDURE

We used a double episcotister in which one disk possessed a single open sector, and the second disk possessed either two or three open sectors. Three different *PCF*'s were used, but in making any of the comparisons between *CFF*'s for the one, two, and three sector disks, a single *PCF* was used at a time. The double episcotister enabled us to make comparisons at any given intensity between the two disks, without stopping to change disks. All that was necessary was to occlude the target controlled by one or the other of the disks. Reducing the time interval between observations under the same controlled external conditions, to reduce the probable variations in unknown factors influencing end-results has obvious merit. Thus, when instrumentation allows this, the procedure is superior to making comparisons between readings taken hours or days apart. The assumption of a constant organism is necessary and is better realized in this procedure.

The intensity levels used were 0.14, 1.4, 14, 140, and 1400 c/ft.² The targets were approximate rectangles subtending a vertical visual angle of approximately 1° and a horizontal visual angle of about 0.8°. Trial series were always made with low intensity readings first, step-by-step ascending to the maximal intensity level. Disk velocities were varied up and down in alternate readings to obtain *CFF*. Table 1 indicates the onset and decline times for pulses with the particular sized disks and beam widths that had to be intercepted for various selected disk rotation speeds. Figure 1 illustrates the relation of pulse taper to number of open sectors used while holding

cycle length (25 msec.) constant. What was earlier said about the encroachment of the taper upon the duration of the cycle occupied by the pulse at full intensity and the intervals between pulses can be readily seen. Figure 2 shows how, with a *PCF* of $1/6$ for a square-wave sequence, the taper shifts as disk rotation speeds vary.

TABLE 1
THE RELATION BETWEEN DISK REVOLUTIONS AND ONSET AND DECLINE TIMES
OF PHOTIC PULSES

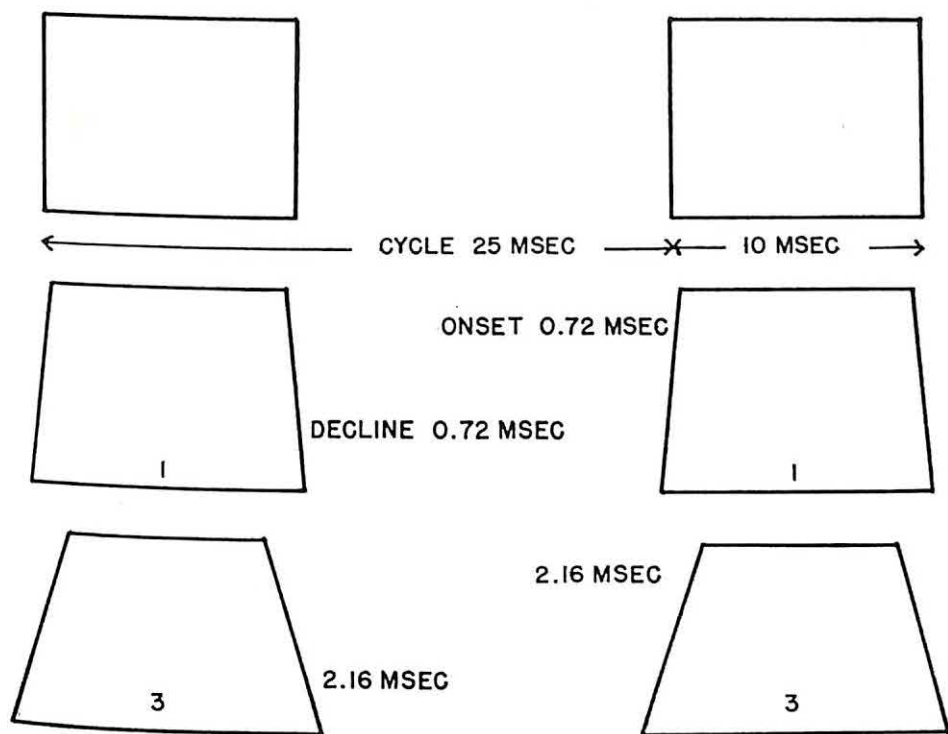
Disk revs (per sec)	Onset and decline time (msec)		
	1 sector	2 sector	3 sector
50	.58	1.16	1.74
45	.64	1.28	1.92
40	.72	1.44	2.16
35	.82	1.64	2.46
30	.96	1.92	2.88
25	1.15	2.30	3.45
20	1.44	2.88	4.32
15	1.92	3.84	5.76
10	2.88	5.76	8.64

C. RESULTS

The results are shown in Table 2 and Figures 3 and 4. It will be noted that the two observers were quite similar in their responses for comparable conditions. Both obtained different *CFF*'s in accord with whether one, two, or three sectored disks were used (with ratio of open to closed sector of disks held constant for each comparison). For both observers, the *CFF* differences in results between the disk with one open sector and the disk with three open sectors was greater than the results when two and one open sectors were used. In all cases, the curves crossed or converged at low levels of intensity. This, of course, meant that, at these levels, the two disks were equivalent in producing *CFF*. The fact that the curves crossed meant that the relative effectiveness of the two disks being compared was reversed as still lower intensity levels were used.

Since it is known that the greater the number of open sectors the more gradual the taper of the onsets and declines of the photic pulses, it can be said that at high intensity levels the more gradual the taper, the higher the *CFF* produced; and at some intermediate intensity the two rates of taper produce the same result. And, finally, at very low levels of intensity, the steeper the taper, the higher the *CFF* produced. It is to be recognized that another factor in addition to rate of pulse taper was introduced by varying the number of open sectors in our disks. This was the actual spacing between pulses. The more gradual the taper, the closer together were the pulses,

since, as was said earlier, their onsets and declines encroached upon the between-pulse intervals. The same encroachment upon the duration of the pulses at full intensity was also brought about. This is to say, only the mid points in the taper pattern of the pulses continued to represent the *PCF*'s indicated for the disks.

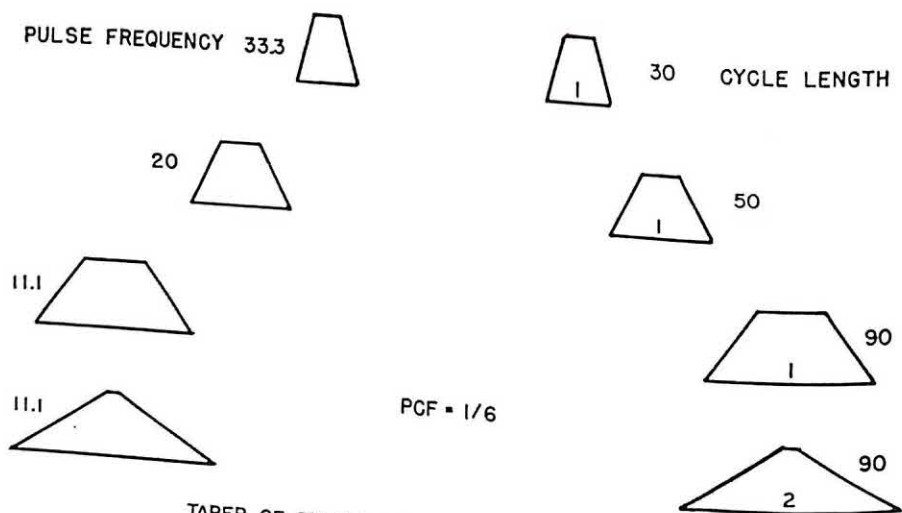


RELATION OF PULSE SHAPE TO NUMBER OF DISK SECTORS

FIGURE 1
DIAGRAMS SHOWING THE RELATION OF TAPER OF ONSET AND DECLINE OF PHOTIC PULSES PROVIDED BY THE EPISCOTISTER ARRANGEMENT USED

This being the case, the durations between the starting and ending points of the gradually tapered pulses occupied greater fractions of the impingement cycles. This altered duration would be more effective for high intensity pulses than for extremely low ones. Threshold considerations dictate this conclusion. For this reason, since at high intensities the three and two sectored disks were more effective than the single sectored disk, the major factor may be thought to be reducible to the change in effective lengths of pulses in the cycles used. To check on this, we would have to examine other

findings in which *PCF* is the variable. Examination of such data (results of a study in press) does not indicate this sort of a relation between *CFF* and *PCF*. Our highest *CFF* values were obtained with *PCF*'s of about 0.25 and generally progressively declined as *PCF* was raised (1). Hence, it is not likely that this factor was responsible for raising our *CFF*'s. It would seem that, in lieu of any other known factor operating here, rate of taper in and of itself is probably the important factor.



TAPER OF PULSE UNDER VARIOUS CONDITIONS

FIGURE 2
DIAGRAMS SHOWING THE TAPER OF PHOTIC PULSES PRODUCED BY EPISCOTISTER UNDER SEVERAL CONDITIONS

On the left, pulse frequencies are indicated, and on the right, cycle lengths. The lower the frequency, the more extended the taper of the pulse in reaching maximum intensity. *PCF* values are measured from zero intensity between pulses to half-intensity of pulses.

D. DISCUSSION

As already mentioned, in the earlier literature on intermittent photic stimulation, the number of sectors in a revolving disk was seen to determine the value of *CFF*. No effective explanation for this was given, although it was recognized that the number of open sectors in relation to the width of beam being interrupted changed the temporal characteristics of the photic pulses used.

Our concern has to do with the relation between pulse taper and *CFF*, and ultimately with the intervening body processes. Thus, one must consider the following: (a) photochemical processes involved in initiating impulses

TABLE 2
RELATIONS BETWEEN PCF, NUMBER OF OPEN SECTORS IN EPISCOTISTER AND INTENSITY:
NUMBERS ARE CFF's IN CYCLES PER SECOND

No. of sectors Intensity	PCF 1/4		PCF 3/4		PCF 1/2		PCF 1/2	
	1	2	1	2	1	2	1	3
<i>Observer N</i>								
.14	22.0	18.7	24.6	18.7	22.0	21.2	21.3	22.6
1.4	26.5	26.1	27.2	26.8	28.6	29.4	29.0	31.2
14.0	38.1	38.0	35.5	38.6	38.8	40.5	41.1	25.1
140.0	46.7	50.6	40.4	41.7	46.3	50.5	47.5	57.4
1400.0	35.4	52.8	37.1	39.0	36.4	39.6	46.1	55.4
<i>Observer B</i>								
.14	21.1	19.9	19.9	18.4	21.0	21.8	21.1	20.7
1.4	26.4	26.6	25.3	25.6	27.6	28.0	27.4	31.2
14.0	36.1	36.2	34.2	35.1	35.5	37.1	37.5	41.1
140.0	41.8	41.8	40.6	42.0	40.6	47.3	43.1	49.6
1400.0	36.1	39.8	34.0	37.6	35.6	39.4	41.9	56.1

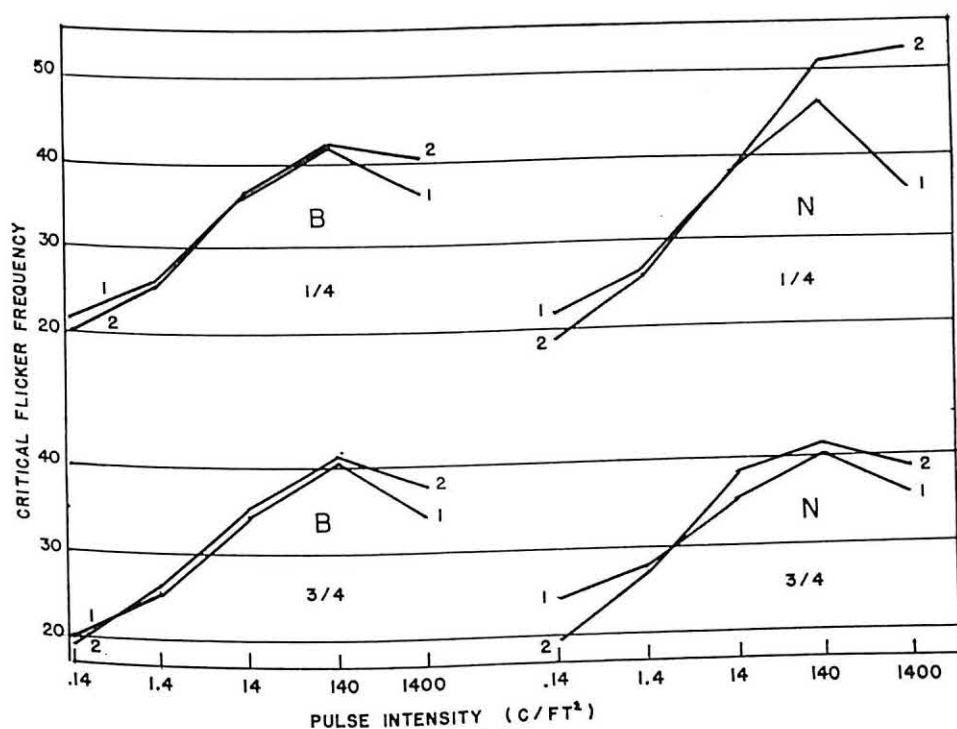


FIGURE 3

CURVES FOR TWO OBSERVERS, N AND B, FOR PCF's = 1/4 AND 3/4, RESPECTIVELY. THE SMALL NUMBERS FOR THE CURVES INDICATE NUMBERS OF OPEN SECTORS IN DISKS USED

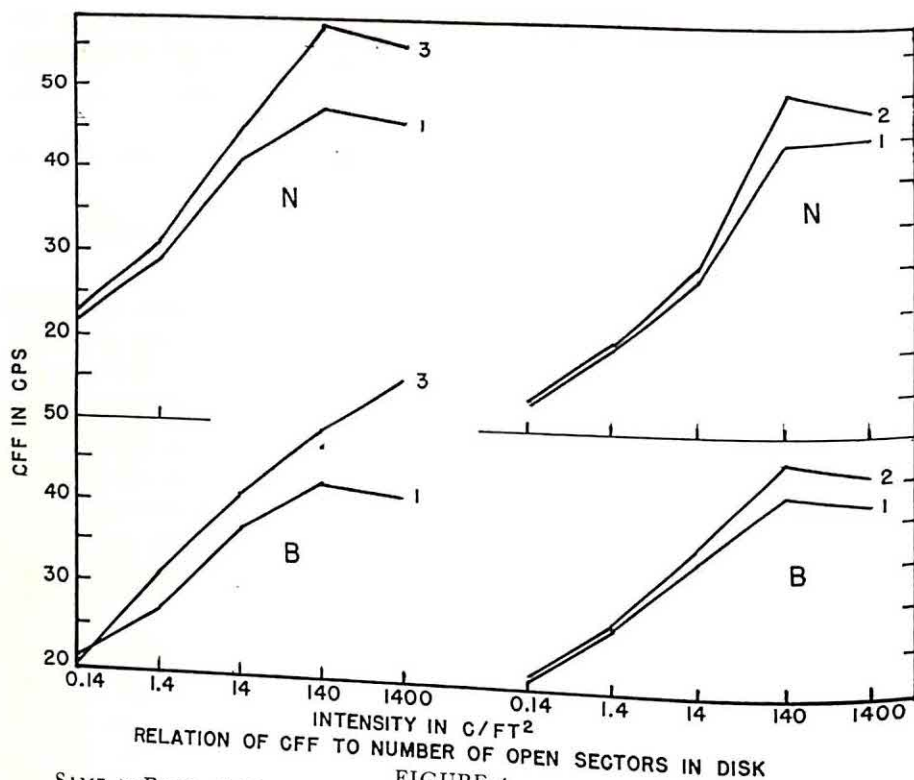


FIGURE 4
SAME AS FIGURE 3, EXCEPT FOR $PCF = 1/2$ AND FOR 1, 2, AND 3 SECTORED DISKS

in receptors, (b) neural processes in the retina, (c) neural processes in the lateral geniculate body, and (d) neural processes in the optic cortex. It has long been customary to begin such an analysis by attempting first to determine the nature of processes at the periphery. The extended work of Hecht and colleagues several decades ago on the photochemistry of retinal receptors is an example of this. A part of what might be said about retinal photochemistry is testable by what we find receptors doing under known conditions. This is to say, much of our knowledge here consists in what we have learned about the discharge of receptor impulses. The psychologists' problem is relating personalistic behavior of the organism to environmental conditions. Sensory processes in this case are personalistic expressions of relation of the organism and its surrounds. Once we know the concomitant events in the sensory and in the stimulus realms, our understandings are furthered by penetrating the intervening realms of body process and mechanism to which we have just alluded. Since the optic pathway is made up of a number of

components whose structure and process differ from one another and the activity of which is dependent upon the prior activity (input) of preceding components, it is not easy to assign responsibility for the contributions each one makes to the final sensory end result.

On this account, there are times and purposes for which the investigational sequence of periphery to center already mentioned must be reversed. This reversal is dependent not only upon the fact just stated, but also upon the fact that the center is itself accessible to study.

But at this point, it does not seem that the proper information about the activities in the various segments of the optic pathway is at hand for us to indicate why the taper of photic pulses is effective in the particular ways we find it to be. All we can say has to do with applying what we know about pulse intensity and the latency of physiological effects, stimulus thresholds, and numbers of receptors activated, etc. It is important, nevertheless, to know the pattern of the effects, i.e., the relations between taper, intensity, and *PCF*. The adequate explanation of these relations awaits further understanding of body processes already indicated. The present findings furnish some fundamental facts to be accounted for.

The information we have obtained provides some material which will be useful in connection with the other phases of any overall study of the relations between photic impingements, photochemical and neural processes, and sensory outcomes. The present work is part of an extensive investigation of ours into such matters and relies for its meaning upon other parts of such a general inquiry.

E. SUMMARY

Critical flicker frequency was determined using an episcotister with disks having one, two, and three open sectors. Not only the number of open sectors, but also the ratio between open and closed sectors of the disks was varied. Ratios of $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$ were used and intensity was varied over a range from .014 c/ft² to 1400 c/ft² in logarithmic steps. It was found that at high intensities the three open-sectored disks required higher repetition rates to produce fusion (*CFF*'s) than the two-sectored disks, and that the two-sectored disks required higher rates than single open-sectored disks. At lower intensities, the curves depicting relations between intensity of photic pulse and *CFF* tended to converge or to cross. This means that the relations just described either reversed or gave some indication of reversing if intensity were reduced below the levels used.

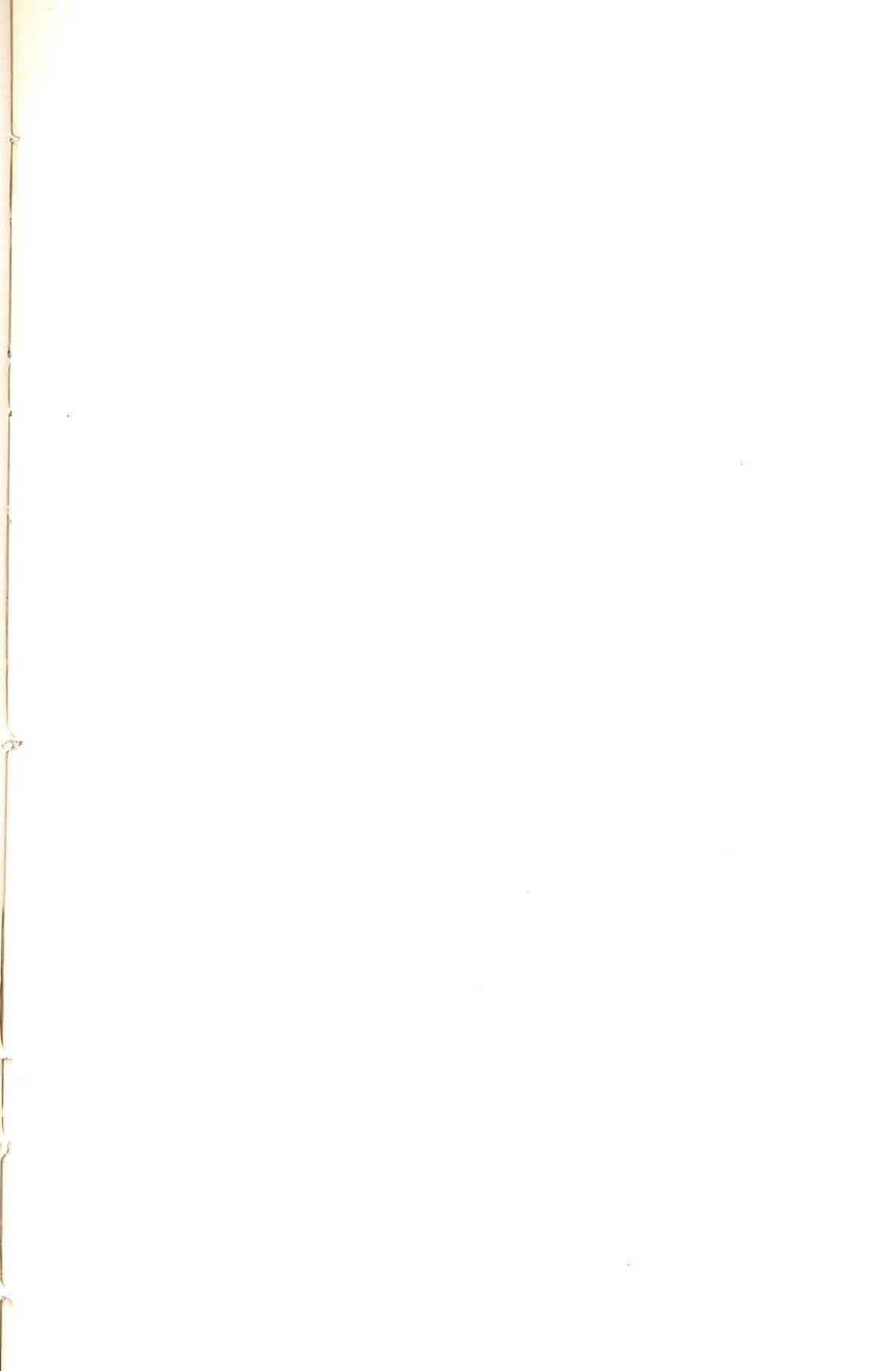
Interpreted in terms of taper of onset and decline of photic pulse, it meant

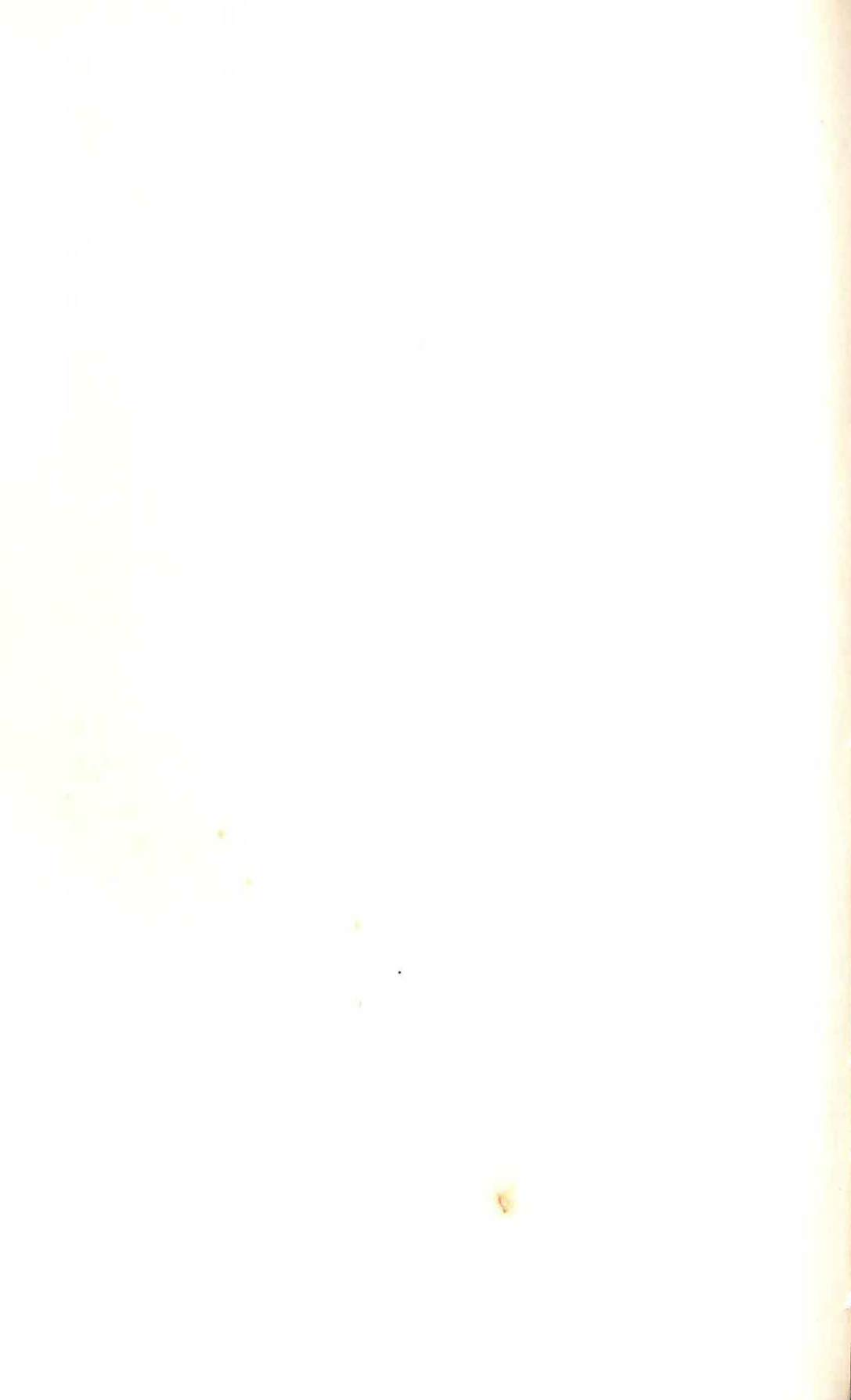
that, at high intensities, the gradually tapered pulses were more effective than the steeply tapered ones in producing high CFF's, and at lower intensities the difference vanished or the relationship between rate of taper and CFF reversed.

REFERENCES

1. BARTLEY, S. H., & NELSON, T. M. Some relations between pulse-to-cycle fraction and critical flicker frequency. *Percept. & Mot. Skills*, (in press).
2. BERGER, C. Illumination of surrounding field and flicker fusion frequency with foveal images of different sizes. *Acta Physiol. Scand.*, 1954, **30**, 161-170.
3. BROWN, C. R., & FORSYTH, D. M. Fusion contour for intermittent stimuli of alternating durations. *Science*, 1959, **129**, 390-391.
4. FILEHNE, W. Ueber den Entstehungsort des Lichtstaubes, der Starrblindheit und der Nachbilder. *Graefes Arch. Ophthal.*, 1885, **31** (abt 2), 1-30.
5. HALSTEAD, W. C. A note on the Bartley effect in the estimation of equivalent brightness. *J. Exp. Psychol.*, 1941, **28**, 524-528.
6. IRELAND, F. H. A comparison of critical flicker frequencies under conditions of monocular and binocular stimulation. *J. Exp. Psychol.*, 1950, **40**, 282-286.
7. LANDIS, C. Effect of varying stimulus conditions on the flicker fusion threshold when nitro-glycerin or the cold pressor test was used. *J. Appl. Physiol.*, 1955, **7**, 513-518.
8. MISIAK, H. The decrease of critical flicker frequency with age. *Science*, 1951, **113**, 551-552.
9. MOWBRAY, G. H., & GEBHARD, J. W. Differential sensitivity of the eye to intermittent white light. *Science*, 1955, **121**, 173-175.
10. SCHENCK, F. Ueber intermittierende Netzhautreizung. XII. *Pflüg. Arch. ges. Physiol.*, 1906, **112**, 292-301.
11. SCHENCK, F., & JUST, W. Ueber intermittierende Netzhautreizung. X. *Pflüg. Arch. ges. Physiol.*, 1902, **90**, 270-284.

Department of Psychology
Michigan State University
East Lansing, Michigan





PARENT-CHILD PSYCHOTHERAPY: A FOLLOW-UP STUDY COMPARING TWO TECHNIQUES*

Psychological Testing Center, Washington, D. C.

E. LAKIN PHILLIPS

A. PROBLEM

Studies of the results of psychotherapy with children are not numerous. It is estimated that between 80 and 90 per cent of published reports on child therapy are on interesting cases and syndromes; few studies exist comparing two techniques of therapy in the same setting.

Levitt (10) surveyed a number of reports on results from child therapy and failed to find clear support for the view that psychotherapy with neurotic children is effective. His study is reminiscent of Eysenck's similar review of results on adult psychotherapy (13).

Other follow-up studies (1-4, 6-9, 11, 16-20) of the results of child psychotherapy range from 50-75 per cent reported improvement (with the remainder reporting failure or no improvement), except for the Phillips and Johnston study showing an 87 per cent improvement for "time limited" therapy as against 57 per cent improvement rate for "conventional" therapy in two child guidance centers. It is fair to say that the median figure representing improvement, in a two-way breakdown of results, is 65 per cent in the vast majority of studies. If this figure is inserted in a 2 x 2 contingency table, it turns out that a clinic has to treat about 100 children in order to have this level of success allow results that are reliably better than chance.

Pitted against this success ratio and its vicissitudes is a hidden cost: the large attrition rate. It is not unlikely that the high drop-out rate among child as well as adult cases, is great enough to warrant saying that a clinic would have to diagnose and process 150-200 cases before arriving at 100 therapy cases, of which 65 per cent would be expected to show varying degrees of improvement.

B. PROCEDURE

Over several years the writer (11-15) has experimented with methods of overcoming the large attrition rate and in improving the success rate from parent-child psychotherapy. This study reports on some of the larger and

* Received in the Editorial Office on December 30, 1959, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

grosser aspects of the method of conducting a non-depth therapy, as well as comparing results from this method with those of traditional practice in child guidance clinics.

Very briefly outlined, the method of therapy employed consists in the following emphasis:

1. Therapy begins with the first interview; no separate diagnostic function exists. Whether therapy will continue depends upon agreement between patient and therapist.
2. Therapy hours are divided about equally between parent and child. This cuts the interview time per case exactly 50 per cent.
3. Except for the first two or three interviews, meetings are held twice a month or less often. This saves time, also.
4. Unless permission is denied the therapist, he contacts the child's teacher to get a description of problems at school; suggestions are given and received from the teacher; so that a *continuity* of method and outlook is established immediately as to how the child can be constructively handled.
5. The parent is given "common-sense" explanations of what the child's behavior means ("He's trying to get by with this and he probably thinks you won't stop him." "He's had luck with his approach before, so why won't he try it again?" "If you become more definite and consistent, this will help the child know where he stands with you.") Such statements have a salutary effect; they enable the parent to see how the child evidently views his relationships and they aid the parent in taking practical, constructive steps.
6. Stress is placed on "keeping structure," that is, on setting limits on behavior and specific aims for achievement. This allows the parent and child to develop a success pattern in daily relationships rather than being bogged down with failures and impasses.
7. After the initial structuring interviews, parents are encouraged to go on their own for varying periods of time, then check back with a more specific set of successes and failures in coping with the child. In this way a series of only 10 interviews might be distributed over five or six months. A similar contact is maintained with the teacher, the frequency of which depends upon the child's problems and the teacher's ability to structure relationships at school.
8. Homework, household chores, and a fairly set (but *not* rigid) routine is emphasized. Children misbehave, it is hypothesized, largely out of having a too loose structure of requirements and relationships. If this structure is firmed up in sensible, fair, consistent ways, the child will usually improve.

9. Regardless of the degree of disturbance, these general methods are employed; they vary, of course, with differences in children and parents, but this does not alter the general principles.

10. These methods open up therapy to any interested observer. The judgment of success pivots on three bases: (a) Are the symptoms less frequent than before? (b) Are the symptoms less intense (or of less duration) than before? (c) Does the child (and parent) recover faster after any occurrence of upset? Feelings and personal reports are certainly not eschewed, but the real heart of change that is to take place must rest upon *behavioral* grounds.

11. In a general way, the therapy is characterized by a widespread, alert set of relationships with people important to the child, rather than an intensive, personally involved, long-lasting therapy, with the child alone.

C. SAMPLING

Three studies are included in the sampling. Two of these are set in a suburban child guidance clinic, and one in the private practice of the writer.

1. *Study One*

A follow-up questionnaire composed of six quantitatively rated items, using a 7-point rating scale, was sent to the parents of 29 structured (non-depth) cases, with 27 returning the questionnaire; and 48 depth cases, with 30 cases returning the questionnaire; all from a child guidance case load.

Statistical data include: *Age* (Non-depth: Mean 8.2; *SD* 3.1; range 3.5 to 11.4 years; Depth: Mean 8.1; *SD* 2.1; range 3.6 to 12.0 years). *IQ* (Non-depth: Mean 109.1; *SD* 9.27; range 100-135; Depth: Mean 110.0; *SD* 13.80; range 95-140). *Sex* (Non-depth: Boys 70 per cent; Depth: Boys 65 per cent).

2. *Study Two*

Two parent discussion groups ($N = 15$) were given follow-up questionnaire, all of whom returned it. No children seen in therapy. Parents met for 75 minutes one time per month for nine meetings throughout a school year. All cases handled from a non-depth standpoint.

Statistical data include: *Age* (Mean 9.10; *SD* 4.9; range 2.6 to 12.6). *Sex* (9 Boys, 3 Girls). *IQ* (No precise test results available, but all were considered average or above by school psychologists and teachers). Average number of meetings attended by group was 4.40; range was 3-7 meetings.

3. Study Three

This study covers 52 non-depth cases seen in private practice. Statistical data include: *Age* (Mean 9.8; *SD* 4.9; range 2.6 to 12.6). *IQ* (Mean 102.25; *SD* 10.90; range 90-150). *Sex* [Boys, 36 (69 per cent)].

D. RESULTS

Preliminary to the discussion of the findings associated with the above studies, it is necessary to state some related statistical findings.

Since Study One covered cases seen over a previous five-year period, it was important to determine if the lapse of time before the ratings were obtained had any systematic effect. Time-since-therapy-ended was divided into 10 6-months intervals which, in turn, were correlated with ratings on the questionnaire. Two ratings from the questionnaire were analyzed; The therapists and the parents each rated how well the child had improved in an over-all manner; the ratings were independently made.

For 57 cases (30 depth and 27 non-depth), time elapsing since terminal therapy visit, and the therapist's rating of success as per item one, $r = -.04$. Time elapsing since terminal visit and parental ratings of over-all success, $r = .004$. Thus, no significance can apparently be laid to the time factor in influencing ratings by parent or therapist.

The number of visits to the clinic correlated with the therapist's rating of over-all success, $r = .004$. The number of visits and parental rating was, $r = -.28$. The longer the therapy continued, the greater the tendency for the parents to rate the results as less successful.

Table 1 summarizes the results on the number of therapy visits for the depth and non-depth populations, as well as comparing the parental ratings on the six questionnaire items for the two contrasting therapies.

We see from this table that the non-depth results are placed consistently ahead of the depth therapy results by the parental ratings. In the depth group the results were reliably below those of the non-depth cases in the clinic setting, as well as in the parent group meeting (clinic setting) and the 52 cases from private practice.

If we look at the average number of interviews needed to achieve these results, we observe another advantage associated with the structured (non-depth) practices. Generally speaking, it took less than one-half as many interviews to produce the better ratings associated with the structured therapy.

Three other results seem to characterize the structured therapy results: First, each interview period of 45-50 minutes was divided between parent and

child, thus consuming half the time per interview of the depth therapy. Altogether, this means that the non-depth therapists spent less than one-fourth as much time with patients (parent and child) as the depth therapists. Second, the non-depth cases were saved the 5-10 hours of diagnostic workup and preliminary conferences so costly in time and energy. Third, the greatest advantage was that parent and child, in the structured therapy practice, were taken into treatment immediately upon application. There was no long wait first for diagnosis, then for therapy. It is not unlikely that the large attrition rate associated with conventional depth therapy stems in part from the almost interminable wait for help. However, these practices seem indigenous to the depth position and it is not unlikely that if the intake and diagnostic features were modified this would bring a modification of other aspects of depth therapy and practice.

TABLE 1
NUMBER OF INTERVIEWS AND RATINGS FROM QUESTIONNAIRE, ONE DEPTH, AND
THREE NON-DEPTH POPULATIONS

Population	No. interviews*	Questionnaire items**					
		1***	2	3	4	5	6
Depth (N = 30)	17.53 Mean (11.80 SD)	4.27 (1.38)	4.47 (1.31)	4.50 (1.44)	4.65 (1.41)	4.35 (1.33)	4.28 (1.51)
Non-Depth (N = 27) Clinic, Indiv.	7.04 Mean (3.82 SD)	4.89 (1.02)	4.75 (1.15)	4.72 (1.06)	5.09 (1.07)	4.50 (1.30)	4.72 (1.40)
Non-Depth (N = 15) Clinic, Group	4.40 Mean (1.08 SD)	5.27 (1.10)	5.40 (1.11)	4.60 (0.81)	5.08 (0.73)	4.20 (1.06)	4.70 (0.91)
Non-Depth (N = 52) Private, Indiv.	8.02 Mean (4.82 SD)	5.11 (1.26)	5.36 (1.34)	5.12 (1.20)	5.56 (1.62)	4.83 (1.13)	4.73 (1.00)

* Differences between the mean number of interviews between depth and non-depth populations are all reliable beyond the .01 level.

** Differences between the means of each of the six questionnaire items, between the depth and each non-depth population, are reliable beyond the .05 level.

*** Questionnaire items, 1-6, inclusively, refer to: improvement in original complaint; parent's ability to handle child better; child's behavior at home; child's behavior in formal groups; child's behavior in informal play groups; behavior specifically at school.

An additional interesting finding calls for some attention. In the first study, 30 per cent of the 180 ratings by the depth population parents (6 questionnaire items times 30 parental ratings) were rated "6" or "7," the two highest ratings. In the structured therapy population, the comparable level occurred in 29.1 per cent of the time. Thus, with both therapies the highest ratings were comparable in frequency and proportion. However, throughout the whole rating range, from failure through a state of no change to modest changes, the non-depth procedures yielded higher percentages of success. For

example, parental ratings on item number one (as to general success in therapy) yielded 25 successful ratings out of 27 cases (92 per cent) in the non-depth population; while the parents of the depth cases rated 18 cases out of 30 as successful (60 per cent), from among 48 sent questionnaires. Comparably, in the second study, 14 of the 15 non-depth treated parents (93 per cent) judged the therapy as successful; and in the third study, 48 of the 52 privately treated cases judged the therapy as successful (92 per cent). In 12 instances in the latter group the ratings were received within three months after termination of treatment because of families moving away; the follow-up period for the total privately treated cases averaged one year.

One final set of figures may be illustrative of the relative efficiency of the structured therapy method. During a given span of time—the last three months of the 1956-57 school year—one non-depth therapist had 61 active parent-child cases; three local suburban child guidance clinics, with staffs of 4-6 full time personnel and comparable clerical staffs, had, respectively 12, 30 (but in addition had 50-60 in seven group therapy constellations), and 50 individual therapy cases.

E. INTERPRETATION

Although this is an extremely small and modest effort among all the resourceful therapies extant today, there appears to be some value associated with the simple, matter-of-fact approach to parent-child problems (13). If the non-depth procedures were adopted widely in child guidance and school clinics, it is estimated that from four to five times as many patients could be handled as are now handled by prevailing depth practices. This statement is not made for propaganda purposes but simply to call attention to the necessity of providing wider services for increasingly larger numbers of people.

Depth therapists might protest that these results show brief therapy works simply because in a short time "deep" therapy cannot occur. The reply to this is simply to say that regardless of any circularity between short term therapy and non-depth practices, the results indicate that the briefer procedures work; the reasons for this success and the full measure of implications are yet to be hammered out in clear terms. Any deliberate effort to make therapy shorter may be open to criticism. If we set out to shorten therapy and are successful, what do we show? We may show, among other things, that the routes to improved mental health are multiple and that we can exercise some choice in how we proceed. If we can exercise a choice in therapeutic procedure that both shortens therapy and does not jeopardize results, then we

have improved upon a useful tool; such a fact needs broad recognition and objective evaluation.

Even if no differences in the success-failure ratio were demonstrated between depth therapy and non-depth therapy (or any other contrasting therapies), there might still be a practical advantage to one over the other because of the greater social usefulness of one. Not many psychologists are satisfied with methods of estimating the outcome of therapy, so if an emphasis were shifted from measured outcome to actual services employed, to the hours of professional time and skill needed, and to the applicability of given methods to a wider section of the population needing help, these relatively objective variables could carry the burden of proof of the need for more practical methods and heuristic procedures.

REFERENCES

1. ANDERSON, F. N. Six years of child guidance. *J. Juven. Res.*, 1931, **15**, 73-96.
2. BARBOUR, B. F., & BEDELL, C. J. The follow up of a child guidance clinic program. *J. Ment. Sci.*, 1955, **101**, 794-809.
3. BENJAMIN, A., & WEATHERLY, H. E. Hospital treatment of emotionally disturbed children. *Amer. J. Orthopsychiat.*, 1947, **17**, 665-674.
4. COX, F. N. Socio-metric status and individual adjustment before and after play therapy. *J. Abn. & Soc. Psychol.*, 1953, **48**, 354-356.
5. EYSENCK, J. H. The effects of psychotherapy—an evaluation. *J. Consult. Psychol.*, 1952, **16**, 319-324.
6. FAMILY WELFARE ASSOCIATION OF EVANSTON. We check up on ourselves—A study of 99 cases one year after contact. Evanston, Ill.: Family Service Assn., May, 1939.
7. GIBBS, J. M. Group play therapy. *Brit. J. Med. Psychol.*, 1945, **20**, 244-254.
8. HARCASTLE, D. H. A follow-up study of 100 cases made for the Department of Psychological Medicine, Guy's Hospital. *J. Ment. Sci.*, 1934, **90**, 536-549.
9. HEALY, W., & BRONNER, A. F. Treatment and what happened afterwards. Judge Baker Guidance Center, Boston, Mass., 1939.
10. LEVITT, E. E. The results of psychotherapy with children. *J. Consult. Psychol.*, 1957, **21**, 189-196.
11. PHILLIPS, E. L., & JOHNSTON, M. S. H. Theory and development of short-term, parent-child psychotherapy. *Psychiatry*, 1954, **17**, 267-275.
12. PHILLIPS, E. L. *Psychotherapy: A Modern Theory and Practice*. Englewood Cliffs, N. J.: Prentice-Hall, 1956.
13. ———. Cultural vs introspsychic factors in childhood behavior problem referrals. *J. Clin. Psychol.*, 1956, **12**, 400-401.
14. ———. The use of the teacher as adjunct therapist in child guidance. *Psychiatry*, 1957, **20**, 407-410.
15. ———. Contributions to a learning theory account of childhood autism. *J. of Psychol.*, 1957, **43**, 117-124.
16. ROSENTHAL, F. M., & PINSKY, G. D. Follow-up method in child guidance work. *Amer. J. Orthopsychiat.*, 1936, **6**, 609-615.
17. SEEMAN, J., BARRY, E., & ELLINWOOD, C. Process and outcomes of play therapy. *Amer. Psychol.*, 1956, **11**, 428 (Abstract).

18. WITMER, H. L. A comparison of treatment results in various types of child guidance clinics. *Amer. J. Orthopsychiat.*, 1935, **5**, 351-360.
19. WITMER, H. L., *et al.* The outcome of treatment in a child guidance clinic—a comparison and an evaluation. *Smith Col. Stud. Soc. Work*, 1933, **3**, 333-339.
20. WOLFE, B. M. The later adjustment of sixteen children diagnosed as psychopathic personalities. *Smith Col. Stud. Soc. Work*, 1942, **13**, 156 (Abstract).

Psychological Testing Center

1912 R Street, N.W.

Washington 9, D. C.

CUTANEOUS COMMUNICATION: ABSOLUTE IDENTIFICATION OF ELECTRICAL INTENSITY LEVEL*

Psychology Division, U. S. Army Medical Research Laboratory

GLENN R. HAWKES

A. INTRODUCTION

The skin has often been considered a candidate for communication when the need arose to find a substitute for the eye or the ear. The Braille system, of course, has made a vast array of literature available to the blind, but long training is required to master this system, and the delay involved in encoding material into Braille severely restricts information transmission speed.

Many attempts have been made to train *O*s to "hear" through the skin by applying to the fingers the vibrations produced by speech (6, 9, 12, 14), but on the whole without much success. While there are similarities between audition and cutaneous sensitivity (4), the basic difficulty with these studies has been the failure to appreciate the fact that the skin cannot make temporal discriminations as fine as those made by the ears. One should not assume that the sensory properties of the skin are so similar to those of the ear that nothing more than a little training is required to enable the immediate interpretation of speech applied directly to the skin. Rather, the first step is to assess cutaneous sensitivity to the types of signals likely to prove useful in communication.

A communication system using mechanical vibration of the skin has been described by Geldard (7). Primary cues utilized were three intensities of stimulation, five loci, and three stimulus durations; combination of these cues afforded some 45 different signals. Training of *O*s with the system demonstrated that these vibratory signals could be used for communication purposes.

An electrical cutaneous communication system analogous to the above would have the advantage of providing a more easily controllable source of stimulation. Also, electrodes are more easily mounted on the skin's surface than are mechanical vibrators, and are typically less bulky, thus affording less interference with normal movements. Further, for mechanical vibration the amount of power needed to energize the vibrators used was

* Received in the Editorial Office on January 6, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

on the order of 10 W., as compared to only a tenth as much (8), or less, necessary for electrical stimulation of the skin.

In designing an electrical cutaneous communication system, investigations may profitably be made of those cues already demonstrated to be useful in the mechanical vibration system, viz. intensity, duration, and locus of stimulation. A report on the discrimination limits for frequency and intensity of electrical currents has been made (2) with the suggestion that this type of cutaneous stimulation might prove useful for communication. Further, information is available on the size of ΔI for several frequencies and standard intensity levels (10). The size of ΔI indicates, however, discrimination limits under laboratory conditions. Communication systems require the O to identify a number of signals transmitted to him at relatively rapid rates. The ability of the O to make absolute identifications of intensity levels must be investigated in order to select signals which will minimize confusion and errors.

Alluisi (1) has listed four factors which determine efficiency of absolute identifications: (a) the number of response categories in relation to the number of stimulus categories; (b) the spacing along the dimension in question of the stimuli to be identified; (c) knowledge of results; and (d) the extent of the physical range of stimulation used. In the study of absolute identification of vibratory intensity on which Geldard's system was based (15), the number of response categories was equal to the number of stimulus categories, a situation considered optimal by Alluisi. Subjective equality of spacing of the stimuli used also was said to be optimal; in the mechanical vibration study the stimuli were an equal number of *j.n.d.*'s apart along the intensity dimension. This spacing procedure assumes equality in size of the *j.n.d.*'s, an assumption implicit in Fechner's formulation of psychophysical relationships. Investigations of a number of sensory continua by Stevens (16, 17) have demonstrated that *j.n.d.*'s are not necessarily subjectively equal.

The method of "magnitude estimation" has been the one principally used by Stevens to examine sensory continua in an effort to relate them meaningfully to physical variations such as changes of intensity of stimulation. Data on most of these continua have yielded functions which, plotted in log-log coordinates, may be fitted by a straight line whose slope is the exponent of a power function relating the physical dimension to apparent magnitude of the sensation. The slope of the line which best fitted such data for intensity of electrical stimulation of the skin has been reported to be 3.5 (11, 18). Hawkes' results are reproduced in Figure 1.

The number of possible absolute identifications of current intensity levels

may be examined in regard to both the percentage of correct identifications and the amount of information transmitted by the *O*s. Information theory, considering the *O* as a communication channel, predicts that an increase in the amount of input information results in an increase in information transmitted that eventually levels off at some asymptotic value. Miller (13) has suggested that the limit for undimensional absolute judgments appears to be about seven (about 2.6 bits). Investigations of the number of absolute identifications when the stimulus is mechanical vibration of the skin have indicated that the channel capacity for this kind of stimulation may be less, perhaps only three (7).

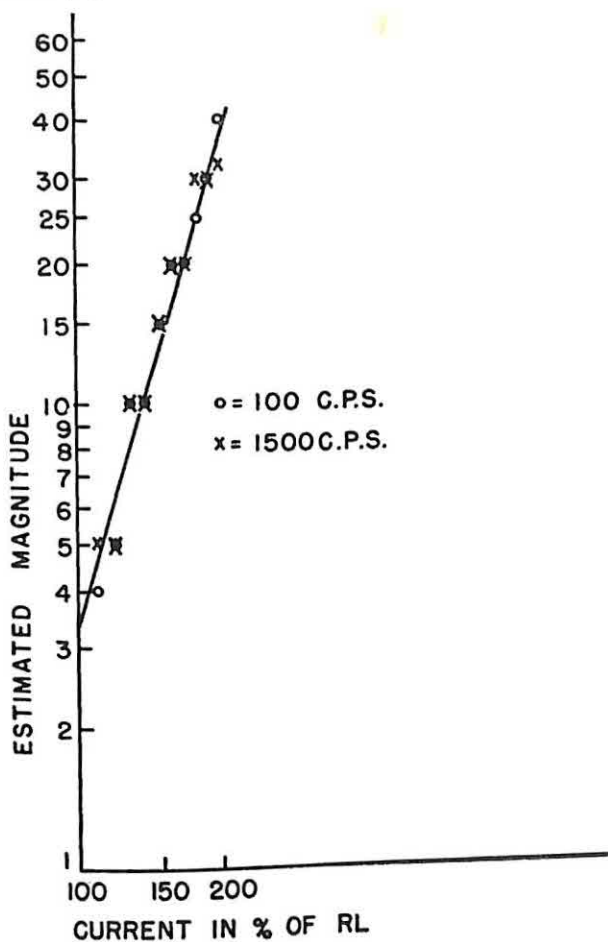


FIGURE 1
EFFECT OF CURRENT INTENSITY ON ESTIMATED MAGNITUDE OF SENSATION FOR TWO
FREQUENCIES OF ALTERNATING CURRENT

The present study was designed to determine the number of current intensity levels which could be absolutely identified with 100 per cent accuracy, and the channel capacity for such stimuli. Knowledge of the number of intensity levels 100 per cent correctly identified would be useful in the design of an electrical cutaneous communication system requiring perfect accuracy with minimal training, while knowledge of the number of levels which achieves maximal transmission of information would be useful in the design of a system which could tolerate some error. The intensity levels used in this study were equally spaced in terms of apparent subjective magnitude of sensation according to the results shown in Figure 1, for from the evidence accumulated by Stevens, it appears that a scale of subjective intensity would be more closely approximated by the direct estimation method than by the use of *j.n.d.*'s.

B. METHOD

Twenty-four *O*s participated in this experiment. Eighteen had no previous experience in the reporting of sensations elicited by the application of alternating current stimulation to the skin. This group was designated as "naïve." The remaining six subjects were experimental psychologists, and had participated in many psychological experiments. This group was designated as "sophisticated."

A block circuit diagram of the apparatus is shown in Figure 2. The electronic switch was designed to turn the signal on and off at rates of onset and offset which avoided the production of transients. Current flow was computed from the voltage change across a 100-ohm resistor in series with the subject. Also in series with the subject during the experiment proper was a 500,000-ohm resistor to reduce fluctuations in the amount of current flow due to the changes in the skin resistance of the subject. The wave form remained sinusoidal for the frequencies and intensities used in this experiment. One lead to the *O* was attached to a 12 mm. diameter circular electrode, coated with electrode paste, which rested on the pad of the index finger of the *O*; the other lead was attached to a 25 mm. diameter circular electrode which rested on the palm.

Four sets of stimuli were selected for absolute identification of current intensity. The two-category set consisted of a low intensity level, just enough higher than the *RL* of the *O* to assure that it would always remain above the *RL* for a series of judgments, and a high intensity level which was below the pain *RL* for most of the subjects; the values selected were 114 and 197 per cent of the *RL* current value. Stimuli for the remaining sets of intensity

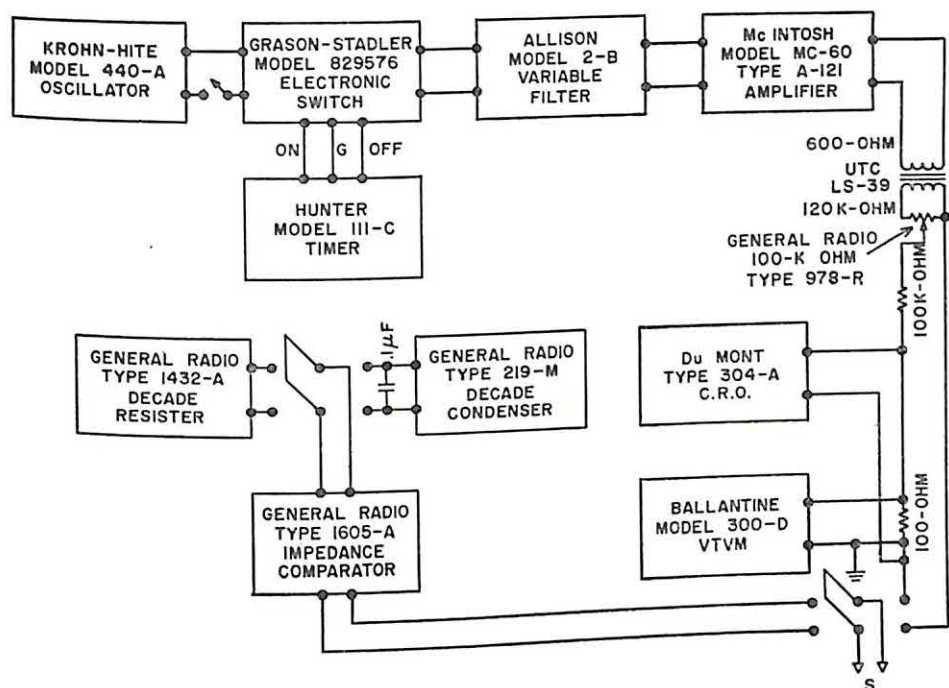


FIGURE 2
BLOCK CIRCUIT DIAGRAM OF APPARATUS

levels were selected so as to be equally spaced in terms of equal subjective increments of sensation, according to the results shown in Figure 1.

Twelve "naïve" *O*s were stimulated with alternating current at a frequency of 100 *cps*; 12 additional *O*s received stimulation at 1500 *cps*. Six of the *O*s in the 1500 *cps* group were "naïve" and six were "sophisticated." After preliminary familiarization, each *O* participated in four sessions, with each session being devoted to identification of a different number of intensity levels. Six of the *O*s of the 100 *cps* group, 3 *O*s of the "naïve" 1500 *cps* group, and 3 *O*s of the "sophisticated" 1500 *cps* group, received five intensity levels during the first session, four levels during the second session, etc.; the remaining 12 *O*s received two intensity levels during the first session, three levels during the second session, etc.

After measurement of his *RL* for a given session, the *O* was presented 100 stimuli for identification, the stimuli being in a different random order for each *O* and each session. Before each stimulus presentation, the *O* was warned by "ready," then the stimulus was turned on for 0.5 sec., after which the *O* indicated the number which should be assigned to it. Stimuli were presented at a rate of about one every 2.5 sec.

C. RESULTS

Mean values for the number of correct identifications are listed in Table 1. Perfect accuracy was possible for the *O*s of this study only for two intensity levels.

TABLE 1
MEAN PER CENT CORRECT IDENTIFICATIONS OF CURRENT INTENSITY LEVELS

N	Frequency in <i>cps</i>	Levels in First Session	Number of levels			
			2	3	4	5
6	100	2				
6	100	5	100	87	72	59
			100	89	77	51
"Naïve" Group						
3	1500	2				
3	1500	5	100	96	79	66
			100	95	68	65
"Sophisticated" Group						
3	1500	2				
3	1500	5	99	91	79	68
			100	91	81	64

The information transmitted as a function of number of intensity levels is plotted in Figure 3 for the 100 *cps* group, and in Figure 4 for the 1500 *cps* group. Analysis of information transmitted was made according to a model described by Garner and Hake (5).

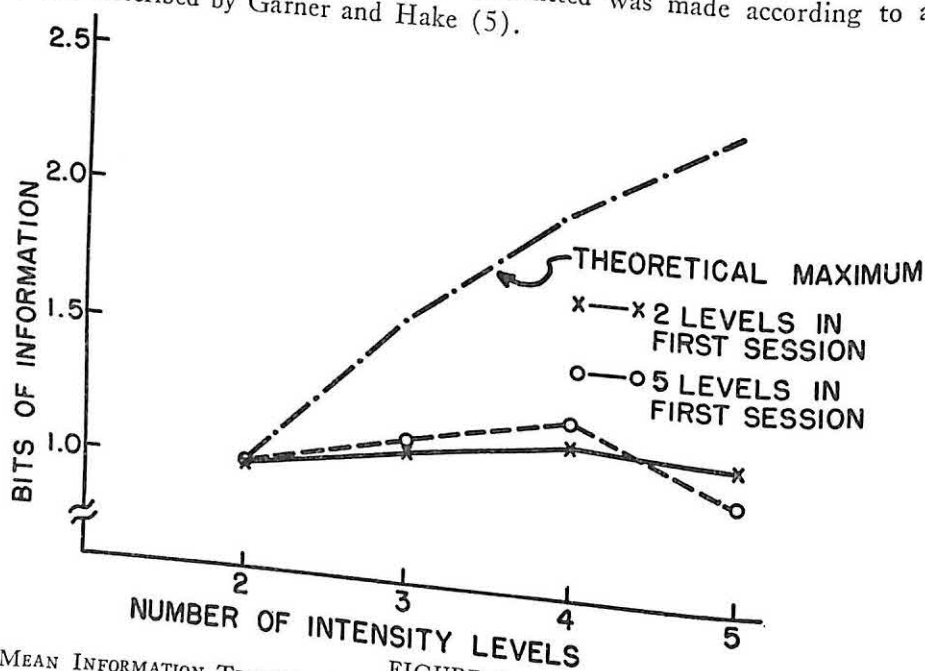


FIGURE 3
MEAN INFORMATION TRANSMITTED AS A FUNCTION OF NUMBER OF CURRENT INTENSITY LEVELS (100 *cps*)

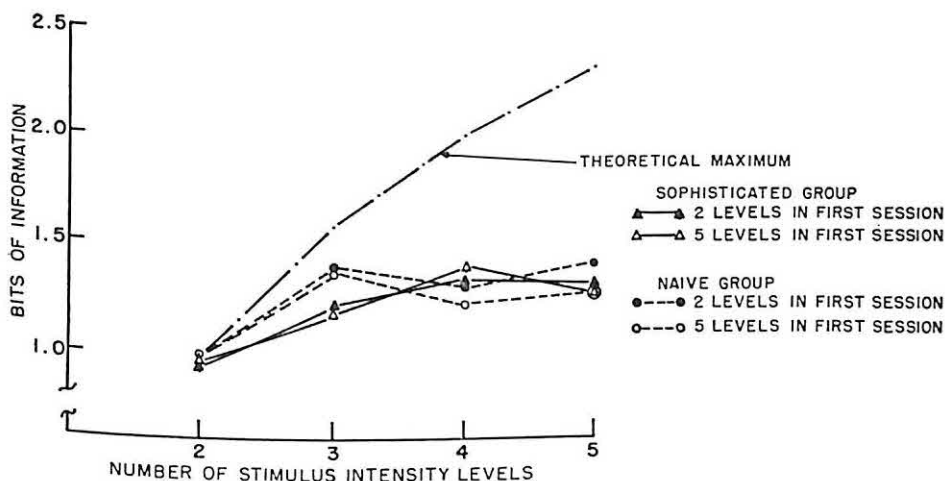


FIGURE 4
MEAN INFORMATION TRANSMITTED AS A FUNCTION OF NUMBER OF CURRENT INTENSITY LEVELS (1500 cps)

A *t*-test for differences in the "amount of information transmitted by the average *O*" (1) indicated that there was no significant difference between the 100 cps group which was presented two intensity levels in the first session and those presented five levels in the first session ($t = .14$). Analysis of variance was made of the differences in information transmitted by the subjects of the 1500 cps group; the results are summarized in Table 2. None of the main effects or interactions approached significance.

A *t*-test was made of the differences between the "amount of information transmitted by the average *O*" for the 100 cps group and the group which received stimulation at 1500 cps. The *t* value of 2.99 was significant at the .01 level of confidence, demonstrating a significantly greater amount of in-

TABLE 2
ANALYSIS OF VARIANCE OF MEAN AMOUNT OF INFORMATION TRANSMITTED BY EACH SUBJECT

Source of variation	df	Mean square	F
Level of sophistication (A)	1	.0051	.32
Order of presentation (B)	1	.0013	.08
Interaction <i>A</i> x <i>B</i>	1	.0038	.24
Within groups	8	.0160	
TOTAL	11		

formation transmitted when the stimulation was presented at 1500 *cps* than when 100 *cps* was used.

D. DISCUSSION

If, in a given situation, 100 per cent correct performance is necessary for communication involving alternating current stimulation of the skin, then only two levels of intensity should be used. From the data, it appears that it is possible, however, to take advantage of the fact that more information is transmitted by three categories than two if intensity level is combined with other cues to add redundancy to a message. Context might also help to reduce the number of incorrect responses due to occasional error if three intensity levels were used in order to achieve maximum transmission of information. It is also possible that with proper training *O*s could identify three levels 100 per cent correctly.

Alluisi has pointed out that immediate knowledge of results may result in an increase in the amount of information transmitted. Such knowledge was not afforded the *O*s of the present study, however, for most communication systems do not permit such feedback. Alluisi also reviewed several investigations in audition and vision which reported a gain in information transmitted as a result of extending the physical range of stimulation. The physical range of the present investigation, however, was restricted by the necessity of avoiding pain sensations for most of the *O*s.

Although the number of intensity levels which could be identified by the *O*s of this experiment was considerably less than the seven which might be expected on the basis of Miller's review, the number found here was very close to that which has been found for a number of dimensions using mechanical vibration of the skin, as well as in other modalities such as taste (3). The small number of intensity levels which could be identified in this study compared to the number of identifications which can be made of auditory loudness, for example, might be due to some extent to the narrow range of physical intensities of current which could be used, when compared to similar ranges for audition.

Intensity of stimulation, if used as a cue in an electrical cutaneous communication system, should be presented to *O*s at a frequency of 1500 *cps* as opposed to 100 *cps*. The "naïve" 1500 *cps* group was not different in any apparent way from the 100 *cps* group, and there was no significant difference in the information transmitted by the "naïve" *O*s as opposed to the "sophisticated" *O*s. Therefore, the significantly greater transmission of information by the 1500 *cps* group must be assumed to be due to the use of this stimulus

frequency, rather than to differences between the *O*s, although some of the performance differential might be so explained.

It remains to be seen what variety of signals is possible when using alternating current stimulation of the skin, but investigations might feasibly be made of the possibility of using duration of stimulation and locus of stimulation as cues in a communication system, as well as such cues as, for example, frequency of stimulation.

E. SUMMARY

Communication by stimulation of the skin has potential usefulness where the eye and ear are not available, as in cases of sensory impairment or when the receiver is in a high noise level situation. Attempts to train subjects to "hear" through the skin by applying to the fingers the vibrations produced by speech have met with only limited success.

A communication system using mechanical vibration of the skin has been shown to be feasible. Geldard described a system of this nature using as primary cues intensity, duration, and locus of stimulation. An electrical cutaneous communication system would have several advantages, e.g., less bulk and power requirement. To design such a system, the ability of the *O* to make absolute identifications of current intensity levels must be investigated.

The present study was designed to determine the number of current intensity levels which could be absolutely identified with perfect accuracy, and the channel capacity for such stimuli. The intensity levels used in this study were equally spaced in terms of apparent subjective magnitude of sensation within the range producing vibratory sensations but not pain.

*O*s could absolutely identify two levels without error, with maximum transmission of information being achieved with use of about three levels. More information was transmitted when alternating current was presented at 1500 *cps* than at 100 *cps*.

Two intensity levels would be useful in a communication system requiring perfect accuracy, while three levels would maximize transmission of information in a system which could tolerate some error.

REFERENCES

1. ALLUISTI, E. A. Conditions affecting the amount of information in absolute judgments. *Psychol. Rev.*, 1957, **64**, 97-103.
2. ANDERSON, A. B., & MUNSON, W. A. Electrical excitation of nerves in the skin at audiofrequencies. *J. Acoust. Soc. Amer.*, 1951, **23**, 155-159.
3. BEEBE-CENTER, J. G., ROGERS, M. S., & O'CONNELL, D. N. Transmission of information about sucrose and saline solutions through the sense of taste. *J. of Psychol.*, 1955, **39**, 157-160.

4. VON BÉKÉSY, G. Similarities between hearing and skin sensations. *Psychol. Rev.*, 1959, **66**, 1-22.
5. GARNER, W. R., & HAKE, H. W. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, **58**, 446-459.
6. GAULT, R. H. Hearing through sense organs of touch and vibration. *J. Franklin Inst.*, 1927, **204**, 329-358.
7. GELDARD, F. A. Adventures in tactile literacy. *Amer. Psychol.*, 1957, **12**, 115-124.
8. GILMER, B. v. H. The sensitivity of the fingers to alternating electrical currents. *Amer. J. Psychol.*, 1937, **49**, 444-449.
9. GUELKE, R. W., & HUYSEN, R. M. J. Development of apparatus for the analysis of sound by the sense of touch. *J. Acoust. Soc. Amer.*, 1959, **31**, 799-809.
10. HAWKES, G. R. Cutaneous electrical delta I. *Amer. Psychol.*, 1959, **14**, 386.
11. ———. Cutaneous discrimination of electrical intensity. Unpubl. doctoral diss., Univ. Virginia, 1959.
12. KATZ, D. The vibratory sense and other lectures. *Univ. Maine Stud.*, 1930, **14**, Second Ser., 90-104.
13. MILLER, G. A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, 1956, **63**, 81-97.
14. RÖSLER, G. Über die Vibrationsempfindung. *Z. Exp. Angewand. Psychol.*, 1957, **4**, 549-602.
15. SPECTOR, P. Cutaneous communications systems utilizing mechanical vibration. Unpubl. doctoral diss., Univ. Virginia, 1954.
16. STEVENS, S. S. The direct estimation of sensory magnitudes: loudness. *Amer. J. Psychol.*, 1956, **69**, 1-25.
17. ———. On the psychophysical law. *Psychol. Rev.*, 1957, **64**, 153-181.
18. STEVENS, S. S., CARTON, A. S., & SHICKMAN, G. M. A scale of apparent intensity of electric shock. *J. Exp. Psychol.*, 1958, **56**, 328-334.

Psychology Division

U. S. Army Medical Research Laboratory
Fort Knox, Kentucky

THE VERBAL POWER TEST OF CONCEPT EQUIVALENTS (VPT)*

Ithaca, New York

E. FRANCESCO

A. INTRODUCTION

In the development of measures of intelligence over the past half century there have been several distinct emphases. On the one hand, there has been considerable attention to concepts of intelligence, to the nature-nurture problem, and to the charting of the crescive and maturing organism. Associated with this has been the focus on unravelling and isolating general and specific factors, the location of primary abilities, and so forth. This attention to charting the abilities and examining developmental characteristics in detail has provided the information necessary to understand the variability of intelligence.

On the other hand, all measures of intelligence ultimately stem from social definitions of intelligence, including performance and success in school, and relative occupational achievement. The criterion of classification that is socially relevant and from which the validation of refined measures of intelligence derives is generally conceived as a complex yet molar quality. The practical job of classifying persons on this overall quality may be far removed from the underlying and more subtle questions of the explicit ways in which people differ. Thus, there has been an increasing emphasis on the development of *efficient* general tests. While recognition is given to the fact that verbal ability is not the sole component of intelligence, practical experience indicates that, in a population with a reported rate of illiteracy of only 3 per cent, it provides an efficient estimate. This report describes the *Verbal Power Test* of Concept Equivalents which, while maintaining the reliability and validity levels required for general screening, was designed to be especially simple for respondents to answer.

B. DEVELOPMENT OF THE VPT

One of the biases that is often noticeable in verbal tests is that as the difficulty of the word included increases, so does the length of the word. This

* Received in the Editorial Office on January 7, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

may reflect the wisdom of languages, which may grow in complexity at least in part through the combination of concepts, but it does not alter the possibility that the component concepts may be ordered in difficulty or frequency of relevant use in communication. The *VPT*, utilizing only short words (four letters), demonstrates implicitly that the component concepts can be selected to provide an ordering on difficulty. It needs to be remarked, however, that finding pairs of concepts that are both of four letters and equivalent is quite another matter, and the *VPT* is based on the recognition of whether pairs of concepts are or are not equivalent.

The development of items for the *VPT* proceeded in the following manner. From an intensive dictionary search a panel of about 250 pairs of matched concepts was located (e.g., neat-tidy). Using small samples of subjects, the panel was reduced to 150 pairs. Three forms were projected utilizing 50 pairs randomly distributed in difficulty in each. For each form 50 additional pairs of mismatched concepts also had to be developed (e.g., foot-shoe). The preferred source for the mismatched concepts was to utilize words taken from the matched pairs of the alternate forms. This did not provide a sufficient number of plausible mismatched pairs, however, and to complete original pool. The resulting three independent forms were pretested with samples of high school freshmen and seniors, and after an extensive item analysis the two working forms of the *VPT* were constructed.

The two forms of the *VPT* were designed not only to incorporate the information from the item analyses, but also an obvious and important distinction between the task of recognizing matched and mismatched concepts. This difference became most dramatically visible in the relative difficulty levels of the pairs of concepts. On the whole, the mismatched pairs of concepts have much lower difficulty levels than the matched pairs. Thus, the *VPT* forms are not symmetric with regard to matched and mismatched pairs, in order to avoid the feeling in respondents of an apparently nonsensical test. It was felt this impression could easily arise when the respondent could appropriately recognize that most of the mismatched pairs were indeed mismatched but could not recognize the matched pairs. Manifestly, such a form would need to be answered by the respondent as dominantly mismatched. For this reason, the *VPT* has 63 matched and 42 mismatched pairs of concepts in each form of 105 pairs.

The selection of items for the forms was guided primarily on the basis of efficiency in discriminating high school freshmen and seniors, and correlation with the total score of the provisional forms. Additional factors needed to

be taken into account, however. For example, some items appeared to penalize some knowledge more than no knowledge; that is, while a group that was selected for a low average score answered a given item at the chance level, a group with a higher average score might actually answer the item incorrectly substantially more often than chance. Since substitute items were available that were answered incorrectly more often by chance by both groups, these were preferred.

A similar problem arose at the point of distinguishing superior persons. While on an empirical level a pair of concepts like "geld-spay" operates as an efficient item over the range of respondents, a small number of superior respondents will insist that while parallelism exists in purpose, the processes involved are not equivalent. In cases of this sort decisions were made on the empirical validity of the item as indicated in the item to score correlations.

C. THE EFFICIENCY OF THE *VPT*¹

An early indication of the efficiency of the *VPT* may be seen in a study of its relationship to the Stanford Achievement Test (scores in grade equivalents).¹ The sample of respondents provided appeared to be badly skewed on the low side of performance, particularly on the basis of clinical assessments of the respondents. For each respondent a clinical opinion was available, based on interview, the *SAT*, and other tests. According to this classification all superior and above average persons were included (3 and 13 respectively), and the remainder of the sample was chosen to favor inclusion of persons classified as average. The consequent sample had a mean score on the *SAT* of 8.0 with a standard deviation of 1.8.

In spite of the fact that an attempt was made to make this sample more characteristic of a normal population by stratifying the analysis sample, the group still must be viewed as one skewed toward the lower end of performance. This should have the effect, in correlational analysis, of attenuating the findings. The difficulty level of the *VPT* is designed to have an effective range that is broad but most relevant for the upper 80 per cent of the adult population. As an indication of the relative skew to the lower end of the distribution in regard to the *VPT*, only one respondent had a score of less than 10 wrong on either form. Such a score would be relatively common among college seniors.

¹ I am grateful to the staff at the New Jersey Reformatory, Bordentown, N. J., who provided data included in this report, and to Dr. Edgar F. Borgatta who provided several suggestions helpful to the development of this test. Specimen copies or copies for experimental use are available through correspondence with the author (E. Francesco, P.O. Box 601, Ithaca, N. Y.).

In the analysis presented here data for the *VPT* are given in terms of performance on matched pairs, mismatched pairs, and total score for each form. Reliability data between forms are indicated in Table 1. The Form

TABLE 1

	1	2	3	4	5	6	7	8
1. Stanford Achievement Test	—							
2. Clinical Assessment of Intelligence		.78						
3. <i>VPT</i> , Form A Matched Pairs			.61					
4. <i>VPT</i> , Form A Mismatched Pairs			.57	.53	.74	.64	.41	.74
5. <i>VPT</i> , Form A Total Score				.47	.66	.59	.39	.70
6. <i>VPT</i> , Form B Matched Pairs				.12	.79	.84	.05	.74
7. <i>VPT</i> , Form B Mismatched Pairs					.67	.12	.84	.55
8. <i>VPT</i> , Form B Total Score						.67	.56	.86
							— .02	.84
								.52

Note: Product moment correlations. No corrections made for attenuation.

A total score correlates with Form B with an $r = .86$. The between forms correlation of matched pairs and mismatched pairs is equally high, being .84 in both cases. The remarkable finding in Table 1 is the apparent total independence of the matched and mismatched portions of the *VPT* (.12 and —.02 within the forms, and .12 and .05 between the forms).

The two validation criteria available here, the *SAT* and the clinical assessment, correlated with an $r = .78$. A dependent relationship exists here, since the *SAT* scores were involved in the clinical assessments. The total score for the *VPT* correlated with the *SAT* with an $r = .74$ for both forms, and it is seen that matched and mismatched pairs have made independent contributions to this correlation.

Correlation of the *VPT* to the clinical assessments is striking at a substantial level, the r 's for the total scores on the two forms being .66 and .70. Again, it is seen that the matched and the mismatched pairs make independent contributions to the correlation.

D. SUMMARY AND CONCLUSION

The *VPT* is composed of 63 matched and 42 mismatched pairs of concepts (four-letter words). The task of recognition of the pairs of concepts as matched and mismatched is demonstrably different, not only in difficulty but also in terms of lack of correlation. Both tasks, however, are reliable areas of performance, and both tasks contribute to the prediction of intelligence as measured by the *SAT* and clinical assessment. The *VPT* may be judged as extremely efficient in its operation, both from the point of view of design and predictive performance.

P.O. Box 601
Ithaca, New York

SCREENING SOLUTIONS TO UPGRADE QUALITY: A NEW APPROACH TO PROBLEM SOLVING UNDER CONDITIONS OF UNCERTAINTY*

Department of Psychology, University of Michigan

NORMAN R. F. MAIER^{1,2}

A. PREFACE

This paper seems to need a preface, and I take the liberty of incorporating one. The objective of the screening principles is to raise the level of problem solving discussions when decisions are needed and further information is not available. It is understood that human beings may engage in faulty thinking but when one cannot prove that such thinking is faulty, how does one prevent it? As a matter of fact even when proof is possible, it fails to deter many because of the human tendency to persist in a point of view.

The four principles presented in this paper are reduced from an original number of nine, each of which had three or more subdivisions, making a total of 38 steps. This reduction was necessary because the initial process was too complex to use. It was found, however, that many solutions could be eliminated by several of the steps and this made possible the present simplified process.

It is hoped that the principles discussed will be viewed in the light of a new direction in the method of working with group problem-solving. It is recognized that there is a need for refining the principles and exploring ways in which they can be tested by research. Since some of the principles may appear irrational from one frame of reference it is necessary to regard the principles as useful in upgrading a group's product rather than in assuring a "correct" or even a "best" decision.

Decision making under conditions of uncertainty according to Luce and Raiffa (1) is the condition in which the outcomes of the various alternatives are unknown. This condition is satisfied by the type of problem under consideration in the present study. However, there is also a difference.

* Received in the Editorial Office on January 11, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This paper was written in conjunction with a research project on problem solving supported by Grant Number M-2704, from the National Institutes of Mental Health, Public Health Service.

² The author is indebted to Professor C. H. Coombs and Dr. L. R. Hoffman for their critical comments and suggestions.

Decision making implies a given number of alternatives, whereas in problem solving the alternatives must be created (3). Thus problem solving involves both choice behavior and the finding or creating of alternatives. Our study, therefore, involves both uncertainty as to outcomes and uncertainty as to the nature and number of possible alternatives. This type of problem is too complex to be handled by the current type of models being tested in connection with the decision making process. The screening process, though less accurate and precise, is a new kind of approach to this complex type of problem.

The potential importance of this approach is not merely one of improving the quality of the product of group thinking but also may lead to the exploration and analysis of the nature of the inference process. Human beings make many inferences, they have hunches and claim to have insights. These processes follow logical principles different from those that are pertinent to proof or deductive thinking. As a result one cannot use standard logic in evaluating the inferences made by living organisms, who by nature go beyond what is known.

B. THE NEED FOR A METHOD TO SCREEN SOLUTIONS

A deficiency in facts in a problem situation does not discourage attempts to find a solution. As a matter of fact such a free condition often leads to an abundance of solutions so that a new problem is presented—that of selecting one from among the many possibilities. Whenever the facts in a situation are inadequate, it follows that a good many of the solutions put forth cannot be logically rejected. If an evaluation is to be made, additional procedures must be developed.

Under such conditions it would be desirable either to obtain more information or to test experimentally the merits of various solutions. However, many of life's problems cannot await the evaluation of solutions by these methods so that, for periods of time, there are stages when people can reasonably disagree regarding the merits and demerits of alternate solutions. All might agree that the solutions will differ in value when more is known, but until then, they cannot agree on the relative merits of various solutions.

In science, conflicting theories are reduced in number and essential refinements introduced as research knowledge increases. The function of crucial experiments is to narrow the field. But even in this instance there often is disagreement over the nature of crucial experiments. This means that what one person regards as proof may be insufficient evidence for another. Thus the amount of observational data needed to convince Darwin

of man's evolution was not convincing to many others. When a man makes an inference he takes a "leap in the dark" and makes some additional suppositions. We might say he "makes up" some new facts so that for him his viewpoint or solution is very plausible. If no one possesses facts to the contrary, his "make believe" facts are acceptable to him, while others may feel cool or even antagonistic toward them.

This disagreement among men regarding the merits of solutions becomes even more marked when we examine some of the day to day problems. Actions have to be taken even when the choice made cannot be justified by the facts or circumstances. These actions or solutions to problems may be viewed either as guesses or as hunches. If we consider guesses to be purely a chance selection, then it follows that people should be of equal ability in choosing solutions that will turn out to be effective. However, the term hunch implies something more than chance—it suggests a personal logic. People differ in their hunches and some have hunches that turn out right more often than chance could permit. If people differ in their ability to solve problems when hunches rather than facts primarily determine the solutions developed, it is possible that this difference represents the contribution of creativity. Our immediate objective, however, is to determine whether solutions can be evaluated by procedures that do not make use of experimentation, logical proof, or probability analyses. In other words, when facts and proof cannot be used to select the best solution from a group of possibilities, what sort of methods or skills can operate to insure better than chance solutions?

The problem is further complicated by the fact that people prefer solutions that they played a part in formulating and this may lead to popular choices rather than ones that eventually may be supported by additional facts. Selecting wise solutions under these conditions introduces numerous subjective factors so that questions of utility, statistical probability, and rational behavior are only a part of the total problem solving process.

At present the hypotheses of man-made machines will be probability predictions rather than hunches. Perhaps this is their greatest virtue. In the meantime it is of interest to determine whether solutions to problems, particularly those involving people, can be screened so as to upgrade the quality of a group discussion outcome.

C. ILLUSTRATIVE CASE (2, p. 39)

In order to describe the screening process it seems best to use a specific problem so that the solutions recommended can be evaluated. In working with this material, the solutions were obtained by dividing groups of 10 to

25 persons into teams of 3 or 4 persons each. The teams were asked to serve as consulting firms and were given the assignment of finding solutions. Procedures for obtaining additional information, such as talking to the personnel in the case, were not acceptable as solutions because these were a means to the final solution. Each team was limited to three action steps. The request for having solutions stated in terms of actions to be taken was included because some groups state solutions in terms of objectives if not carefully instructed.

The case of the *Sewing Machine Operators* was then presented. The situational description was read by the leader, after which two persons in the group were appointed to read the dialogue. The final step was to supply each team with a copy of the above mentioned materials and a copy of the plant lay-out shown in Figure 1. A half hour was allowed for obtaining consensus on three action steps.

1. *The Situation*

For the past six months Bill has been the supervisor of a sewing room of 25 girls in a garment factory. The girls are all union members. Until recently, his main problem was that of getting out enough production. Three months ago, however, all sewing machine operations were changed over from hourly to a team-piece-rate system so that production has become fairly satisfactory. Quality is now the big problem; not only are there too many rejects, but serious complaints are coming in from salesmen in the field. Since the girls are not paid for rejected items, it is difficult to understand why they are not more careful.

Bill has been called in by his department foreman, Mr. Johnson, to discuss the matter.

2. *The Dialogue*

Mr. J: Bill, I want to talk to you again about the kind of work your unit is turning out. What's the matter down there anyway?

Bill: Darn it, I don't know. On the old hourly rate the girls weren't turning out anything, and now on this new group piece-rate a lot of the work they do isn't any good. When I make them do it over they say that I am picking on them.

Mr. J: Steve and Eddie aren't having the trouble with their girls that you do with yours.

Bill: Well, I'm not having trouble with all of mine. There is just this small group of five or six who are the real trouble makers. They all want to be finishers and anything but what they are. I've got them spotted next to my desk where I can keep an eye on them, and I tell them that I won't move them until they learn to coöperate. Even so,

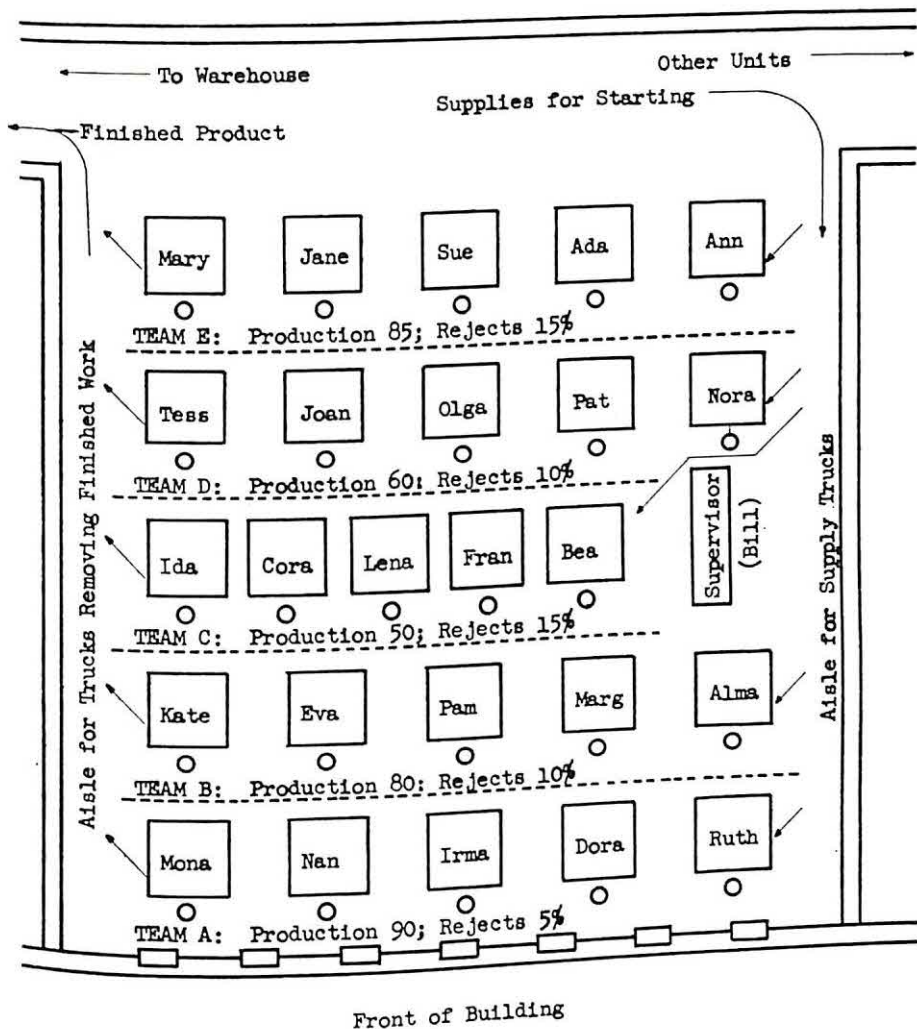


FIGURE 1

SEWING-ROOM LAYOUT

Production Standard, 80 units per team. Quality Standard, 8 per cent rejects.
 Problem Girls: Pat, Nora, Fran, Bea, Marg, and Alma.

I'd like to see Steve or Eddie or anybody else get any work out of them.

Mr. J: You're not trying to tell me that just a few girls out of more than two dozen make your crew look that bad?

Bill: No, but they are the worst ones. I called all the girls in the "C" and "D" teams together last week and gave them a good talking to and now they're worse than ever. Production and quality are both down.

Mr. J: I'm beginning to think that you don't have *any* that are any good.

Bill: No, that's not right. I'll take the girls in my "A" team and put 'em against any we've got. As a matter of fact all of my finishers are a pretty decent bunch. The "B" team has a couple of good workers in it and there is nothing wrong with the "E" team.

Mr. J: Yeah, but their rejects are too high.

Bill: Well, that might be true, but those girls certainly produce. Maybe if I can get them to slow down a little the quality will go up. It's going to make them sore, though.

Mr. J: That's *your* problem. You're not afraid of them, are you?

Bill: No, but they didn't even like it the other day when I got after them for talking on the job. Come to think of it, all that gabbing may be the reason they don't pay any attention to quality.

Mr. J: Well, tell them that if they don't stop talking you'll break up their little club. You're the boss down there, aren't you?

Bill: Well, you've kind of got me there. I hired everybody in Team "E" in one batch with the understanding that they could work together and I hate to go back on my word.

Mr. J: Well, give them a good lecture and threaten to do it.

Bill: I know, but it's a headache and those girls stick together and you can't locate the trouble maker. Even when girls don't get along too well they gang up on you. For example, I gave these girls on Team "C" a safety lecture the other day after one of them got her hands caught, and all they did was gripe and pick on me about everything under the sun. I never saw such a bunch of sour hens before in my life.

Mr. J: What's eating them anyway? Certainly there must have been some one thing they picked on.

Bill: Oh, it was the same old yapping about all of them wanting to be finishers. After they've been on the job a few weeks they think they know everything.

Mr. J: Sounds like you've been giving some of those girls a lot of half-baked ideas about the jobs around here. What's there to being a finisher anyway? The pay is the same.

Bill: I don't know. I think it's just a dumb idea they've got in their heads. You know how women are. I know the end job isn't any easier. *Mr. J:* Was that all they griped about, or do they want us to give 'em the factory too?

Bill: No. Quite a few of them were sore because they said they couldn't make standard. Most of the girls think the bogey is too hard to hit anyway.

Mr. J: I haven't heard any other complaints about it. After all, 80 isn't so high. Why should your girls complain when none of the rest of them do?

Bill: All I know is that they do. Except for Team "A" and a few others who really turn the stuff out, they're just about the worst bunch of eight-balls I ever saw. I don't know why I have to have all of them.

Mr. J: Bill, we've been over all this before and I'm tired of listening to you feel sorry for yourself. Either get those girls on the ball, or we'll have to put somebody down there who knows how to run things. I don't want to be rough about it, but that's the way it is. I'll give you thirty days to get that mess straightened out, and I'll back you up on anything that seems reasonable. If you can show me some results by the end of that time you can stay; if you don't, we'll have to find something else for you. Is that clear?

Bill: I guess so. But after racking my brains like I have for the past six months, I don't know what you or I or anybody else can do with those girls. I've tried everything.

3. *Examples of Solutions Recommended*

The following solutions are typical of those recommended when three solutions are requested. Solutions numbered 1, 2, 4, and 5 are suggested most frequently.

1. Turn lines 90° clockwise (also counter clockwise).
2. Move Bill's desk so as to give Team "C" more space.
3. Discharge Bill.
4. Train Bill in human relations skills.
5. Rotate girls within a team.
6. Rotate positions of teams.
7. Post production records of each team.
8. Obtain individual scores on rejects.
9. Institute music programs.
10. Improve lighting.
11. Put in air conditioning.
12. Put five problem girls in one team.
13. Divide up problem girls and place one or two in each team.
14. Restrict talking in Team "E."
15. Break up Team "E."
16. Wall up back of room.
17. Train Johnson in human relations skills.

4. *Screening Principles*

It will readily be conceded that the 17 solutions listed are excessive in number and vary in excellence. The purpose of the screening process is to reduce this number by eliminating the ones least likely to be successful. In this manner the chances of taking action steps that turn out to be successful should be improved. This improvement in quality should be accomplished without sacrificing the group's acceptance of the selections made.

Two of the screening principles developed are designed to reduce the number of possible solutions by rejecting some of them, and two are designed

to assist in the process of making a positive selection. Solutions that are neither rejected nor selected may be regarded as tabled.

It is necessary to discuss the logic underlying these principles in some detail because it is apparent that the screening process will not conform to the usual logical principles for evaluating a solution. The illustrations taken from the case material are not used to prove the merits of the logic but rather to demonstrate the way each principle can serve its purpose.

a. Negative Principle I. Solutions transferred from other problems should be rejected. Generally speaking it is wise to try to generalize learning, yet this principle states that solutions learned in another situation should not be transferred to this one. The reason for discouraging this type of solution is that it fails to solve a particular problem, but instead solves a class of problems. Acceptable solutions should be tailored to fit the known facts in the case. This does not imply that a solution borrowed from past learning will not be effective. It is granted that such solutions might be highly effective, but they also may be effective in problems other than the one under consideration. Such solutions then become solutions to a type of problem rather than unique remedies for a special condition.³

Although it may be argued that a general solution has greater merit than a specialized one, this insistence on a specialized solution has the following merits: (a) It forces attention on the facts of the case. (b) It rejects pet solutions of group members while eliciting a minimum of defensive behavior.

The mere statement of a problem arouses ideas about what to do to correct the condition. In a group situation this tendency can lead to a considerable waste of time, not only because each person is inclined to make a contribution in the form of a solution, but also because the solutions become the core of the discussion. Thus one solution suggests others and soon the details of the initial problem cease to be examined. In rejecting solutions that are not founded on the facts of the problem, one directs the thinking of the group toward the known facts, even though these be limited in number.

People also have an emotional involvement in their own ideas. To reject a solution because it is a poor one leads to defensive behavior and, because the facts are meager, a good deal of time can be dissipated in face saving discussions. An arbitrary ground rule can screen out solutions in an impersonal manner by classifying them as "out of bounds" even though they

³ It goes without saying that generalized solutions apply to this particular problem when it is shown that the facts upon which the generalization is based are present in this problem. In such cases the facts in the present problem are used to support the solution.

might be regarded as good solutions. Extensive use of this principle has revealed that a person in a discussion group will readily concede that his solution violates a screening principle without becoming defensive.

Recommendations rejected by Negative Principle I are Solutions 7, 8, and 9. Solution 11 also might be suspected to be a borrowed solution and for this reason the persons recommending it should be asked to give their reasons. If their defense is based upon the success with the solution in other offices, then it is of the transfer type and must be rejected; but if the reason given is based upon the arrangement of the windows, it is allowed to be retained. For purposes of the present discussion we will assume it passes Negative Principle I.

Solutions 7, 8, and 9 usually stem from attempts to apply some knowledge about boredom or motivation to solving the problem and hence are of the transfer variety. Invariably persons suggesting them agree that the idea came from other situations.

b. Negative Principle II. Solutions, supported by facts or interpretations of facts that are challenged by other members of the group, should be rejected. It is assumed that each of the solutions not already rejected rests upon facts derived from the problem situation under consideration. Each solution, therefore, rests upon a particular selection or interpretation of facts so that various solutions may have quite different supports. According to this principle, it follows that if a particular set of facts is seen as fictitious by someone else in the group or if a conflicting interpretation can be derived from the facts in the same situation, it is assumed that the solution rests upon an unsound support. Although a complete foundation is too much to expect in the type of problem situation under consideration, it is reasonable to expect some unchallenged interpretations to support the solution. This does not mean that a particular interpretation must be proven sound—rather it must merely be better than those with conflicting interpretations or facts.

The merits of this principle reside in the discussion time saved by eliminating the debate over issues that cancel each other. Interpretations that are opposite and equal in strength not only consume excessive discussion time but also may be regarded as equally weak. The assumption is that high quality solutions would more probably be associated with unchallenged interpretations. It follows therefore that time saved in debating close issues could be well spent in the exploration of other facts and the examination of different interpretations.

Examples of solutions likely to be rejected by Negative Principle II are

Solutions 1, 5, 6, 12, 13, 14, 15, and 16. If Solution 1 is supported by an interpretation that the window position is desirable because Team "A" is the best team, then a conflicting interpretation might be raised with respect to Team "E," which is farthest from the window. If it is argued that all teams should share certain advantages in position, then it is clear that this solution also makes all teams share the disadvantages that are present.

However, if it is argued that turning the lines 90° gives Team "C" more space, this solution should be retained unless it is found that the change takes space from some other team.

Solution 5 usually rests on the interpretation that girls want to be finishers (the last position). The fact, however, is that the girls merely *say* they want to be finishers; perhaps they just want to get away from Bill. However, either interpretation may be used to support the rotation plan. If this is done, then the conflicting interpretation is that rotation within the team places all girls closer to Bill just as much as it places all girls farther from Bill.

Solutions 6, 12, and 13 are subject to similar conflicting interpretations and for this reason are rejected.

Solutions 14, 15, and 16 are based upon the point raised about the talking of the girls in Team "E." The interpretation that talking is undesirable is based upon the poor record of rejects found for Team "E." It could be argued just as reasonably that talking is good, and this interpretation would be supported by the fact that Team "E" has a good production record. It is this conflict in interpretations that eliminates these three solutions.

Solutions 1, 2, 3, 4, 10, 11, and 17 remain after this screening, although 1 and 11 were questioned regarding some of their assumptions.

c. Positive Principle I. Solutions founded either upon (a) any of the unchallenged facts or (b) unchallenged interpretations of facts, taken from the problem situation, should be selected for consideration and evaluation. It also follows that, other things being equal, solutions that have more than one fact or interpretation to support them should be considered to be more appropriate than solutions resting entirely on one of these supports. Furthermore, the more different the supporting facts or interpretations, the greater and more valuable the solution. For example, suppose that research in clinical psychology suggests that a nondirective counseling interview will cause patients to assume responsibility and that research in social psychology indicates that responsibility increases as the leader encourages participation. These two sources of data support the conclusion that participation makes for responsibility. A conclusion supported by such widely different sources of data would be regarded as more valuable than one based upon two related

research studies, such as two studies on interviewing emotionally disturbed people.

We will now turn to the solutions that have not been rejected and determine whether any of them are favored by Positive Principle I. Solutions 2, 3, 4, and 17 can be used to illustrate how this principle selects solutions. Solution 2 has a basic fact to support it: Team "C" has the least space and it is the worst group. Although it can be argued that this relationship might be pure coincidence, such a conclusion does not challenge the fact—it merely is a statement indicating the need for more facts. Furthermore, the interpretation that the observed relationship between the work done and the space occupied by Team "C" is a coincidence, is supported by facts obtained elsewhere and hence violate Negative Principle I. The facts in this situation place the poor production record of Team "C" and its shortage of space together, and it is unlikely that anyone would seriously contend that Team "C" would do worse if it had more space.⁴

Some problem solvers attempt to support Solution 2 with additional facts or interpretations. For example, it may be claimed that the crowding of Team "C" creates problem girls. This interpretation can be challenged by the fact that Teams "B" and "D" have as many problem girls as Team "C," but have the full amount of space. Thus Solution 2 rests primarily on one observed negative relationship between space and productivity.

Solutions 3 and 4 are similar in that they make Bill the problem and as a result they are supported by identical facts. The following observations support these solutions: (a) Steve and Eddie (other supervisors) do not have as much trouble as Bill. (b) The actions that Bill has taken haven't been successful. (c) Finishers are the most satisfied and they are farthest from Bill. (d) The best teams are farthest from Bill.

These interpretations tend to remain unchallenged even though it may be argued that the evidence is not conclusive.

Solution 17 is supported by the fact that Johnson has not been helpful to Bill. This inference usually is not questioned. Sometimes it is argued that training in higher management should precede the training of supervisors. However, this conclusion is based upon facts obtained elsewhere (Negative Principle I) and hence cannot be used to challenge an inference based upon the situation under study. Thus Solution 17 usually rests upon

⁴ It is frequently difficult to convince statistically minded persons to overlook the possibility that errors are produced by a failure to give full recognition to chance. Therefore, we wish to remind the reader that we are dealing with situations lacking adequate facts and no claim is made that Solution 2 is correct. It is merely selected and will have to compete with other solutions.

one support—Johnson's apparent inability to help Bill. If it is claimed that Steve and Eddie work well with Johnson then this lone support is challenged.

d. Positive Principle II. When exceptions to a trend in results can be satisfactorily explained, solutions based upon the trend should be selected for further consideration. Ordinarily an exception to a rule or to a trend in data is regarded as a weakness and hence might result in the rejection of a set of observations. According to Positive Principle II each exception should be examined to determine whether or not it may be associated with some other exceptional condition.

For example, it might be claimed that boys in the fifth grade in school can run faster than boys in the second grade. Let us suppose that in comparing 10 fifth grade boys with 10 second grade boys there are two exceptions to this claim in that two fifth grade boys were beaten in a race by all second grade boys. These exceptions would weaken the above claim even if they did not destroy it. Suppose, however, an examination of the exceptions revealed that the two fifth grade boys who lost the race were cripples. This finding would tend to reinstate confidence in the claim regarding the superior running ability of the fifth grade boys.

When the facts are limited, the appearance of exceptions may cause trends in observational data to be summarily rejected, particularly when no provision is made to give them special attention. Positive Principle II requires that the exceptions themselves must be defended if they are allowed to stand. Any fact in the situation that can explain why the exception occurs, therefore, becomes relevant to this analysis.

Solutions 10 and 11 usually rest upon trends in data. It is found that there is a general tendency for teams nearest the window to do the best. If we take the production figures for Teams *A*, *B*, *C*, *D*, and *E* we find them to be 90, 80, 50, 60, and 85, respectively. Exceptions to a general downward trend are the figures 50 and 85. Thus Team "*C*" does too little and Team "*E*" does too much. The fact that three teams fit the trend, whereas two are exceptions, makes the claim for a trend weak. However, let us examine these

Team "*C*" with the lowest production, is exceptional in another way—it has the least work space. This last fact can explain this team's exceptional performance. The second exception concerns the production of Team "*E*" since this is too high, considering its position. This team also is exceptional in a second way—it is the only team made up of friends. This fact could explain the team's superior production in terms of morale. Granted that nothing is definitely proven, it will be apparent that one exception might eliminate another.

Turning now to reject scores of the various teams, we may examine them in a similar manner. We find the figures for Teams *A*, *B*, *C*, *D*, and *E* to be 5, 10, 15, 10, and 15, respectively. In this instance the rejects of Team *D* may be seen as too low or those of Teams "*C*" and "*E*" as too high. Team "*D*" seems not to have an exceptional condition so cannot be explained away. The high rejects of Team "*C*" can be explained as before, but the high rejects of Team "*E*" now must be explained by some unfavorable factor, such as "working too fast." It is clear that attempts to explain these exceptions away are unconvincing, but it must also be granted that a contradictory inference cannot be drawn.

It is possible to argue that the windows serve a valuable function until the middle of the room is reached, but that beyond this point the distance from them ceases to be a factor. The strength of this argument will vary somewhat from one discussion group to another.

It will be apparent that the case for the trend in results would be greatly strengthened with only minor changes in the data. For the time being let us assume that even though the case for the trend is weak, the suggestion is present. Let us, therefore, see if a distinction can be made between Solutions 10 and 11.

Teams "*C*" and "*E*" have the poorest quality. Since Team "*C*" is poor in all respects this is not surprising, but Team "*E*'s" contradictory performances must be explained. Since Team "*E*" has no problem girls and the production is good, it is more necessary to explain the high rejects than the high production. If this step is agreed upon then the next step can be taken. If this step is not agreed upon, then it follows that the trend in data has been rejected because it is necessary to assume that Team "*E*" is a good team in order to entertain the trend.

Let us assume for the purpose of further analysis that Team "*E*'s" poor quality (high rejects) must be explained. Will poor lighting or poor ventilation explain the poor quality the better? Generally it is agreed that poor lighting accounts for Team "*E*'s" poor quality more effectively than does poor ventilation.

5. *Final Evaluation*

The recommended actions that remain for further consideration are Solutions 2, 3, 4, 10, and 17. Since Solutions 2, 10, and 17 have one support each, while Solutions 3 and 4 have more than one, the latter two actions would be ruled as the ones most likely to produce results. However, Solutions 2 and 10 are relatively inexpensive and might also be acceptable for that reason. Furthermore, Solutions 4 and 17 would be related in that they could

be made a part of a single program. Thus the final decision rests with ways to integrate or combine selected solutions in a practical manner.

In working with the screening principles we have repeatedly been able to cause teams of problem solvers, who were in initial disagreement, to change their selection and settle for Solutions 4 and 10. Solution 3 usually is not accepted because blame is also placed on Johnson, his boss. Since Solutions 3 and 4 are alternatives, preference settles on Solution 4. This degree of agreement cannot be achieved through a general discussion of results. Such discussions drift in various directions and persons tend to defend their initial decisions. Solution 10 rarely survives when the screening principles are not used for selection purposes.

D. DISCUSSION AND SUMMARY

A good deal of discussion time is lost in group problem solving because of difficulties created during the process of trying to solve an initial problem. These difficulties constitute a source of new problems, which interfere with the solving of the original problem because they are purely artifacts created by interpersonal relations. Such factors as injured feelings, faulty communication, personal involvement, insecurity, and face saving can function so as to necessitate compromise solutions or the acceptance of solutions that arouse the least opposition. As a result the products of group thinking often lack objective quality.

This unproductive aspect of problem solving is particularly disturbing when facts are inadequate and freedom of thinking is not restricted by the need to conform to known realities. Often a person will unknowingly advocate a pet solution for a given circumstance because it will solve another problem with which he is faced.

Although good conference leadership can reduce these sources of interference, it is apparent that there is a need for other methods for improving solution quality. Obviously an acceptable evaluation process must be impersonal. The following screening principles were developed to eliminate poor quality solutions and to select those with objective quality. The acceptance of this screening process to persons participating in group problem solving is evidenced by the fact that there is high agreement on merits of the solution that survive the screening process.

The two negative principles that screen out solutions may be stated as follows: (a) solutions transferred from other problems should be rejected; (b) solutions, supported by facts or by interpretations of facts that are challenged by other members of a group, should be rejected.

The two positive principles that select solutions for consideration may be stated as follows: (a) solutions founded either upon any of the unchallenged facts or unchallenged interpretations of facts, taken from the problem situation, should be selected for consideration and evaluation; (b) when exceptions to a trend in results can be satisfactorily explained, solutions based upon the trend should be selected for further consideration.

Solutions that are selected by the positive principles are evaluated by group discussion in terms of (a) cost and practical considerations; (b) the way in which they may be integrated; (c) a selection from among alternatives when two actions are incompatible; (d) an examination of the extent of support given by facts or interpretations; and (e) acceptability to group members.

Solutions that are neither screened out nor selected are tabled and should be considered only when the positive principles have failed to select solutions.

It is apparent that the evaluation process makes use of group discussion, but only after the screening process has forced a thorough consideration of the facts in the situation. Facts cannot be used exclusively with the type of problems considered because they are deficient.

REFERENCES

1. LUCE, R. D., & RAIFFA, H. *Games and Decisions: Introduction and Critical Survey*. New York: Wiley, 1957.
2. MAIER, N. R. F. *Psychology in Industry*. Boston: Houghton Mifflin, 1955.
3. ———. *The Appraisal Interview: Objectives, Methods, and Skills*. New York: Wiley, 1958.

Department of Psychology
University of Michigan
Ann Arbor, Michigan

MODIFICATION OF A COMMONLY USED SOURCE OF ELECTRIC SHOCK*

Department of Psychiatry and Neurology, University of Minnesota Medical School

GUY H. MILES

Several simple devices have been developed for administering electric shock to Ss. One of the most commonly used methods is to place S in the secondary circuit of an isolation transformer. A typical circuit of this kind is shown in Figure 1, where an isolation transformer has been constructed by connecting 2-6.3 v. filament transformers (T_1 and T_2) back-to-back. Shock intensity is regulated by means of an autotransformer (Variac) and shock duration by opening and closing Sw in the transformer circuit. Sw is sometimes connected as shown in Figure 1 and sometimes between S and the secondary transformer winding.

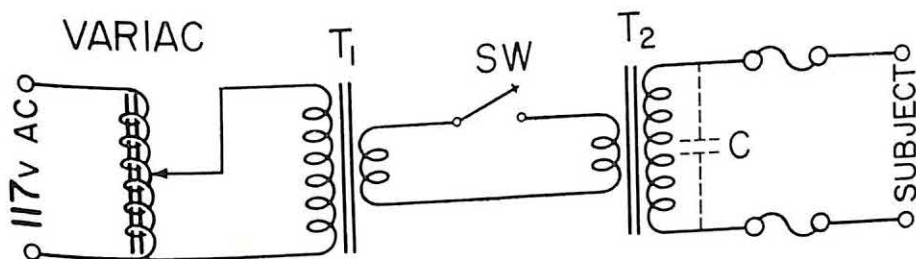


FIGURE 1
CIRCUIT DIAGRAM OF A COMMONLY USED SOURCE OF ELECTRIC SHOCK (SOLID LINES)
AND SUGGESTED SIMPLE MODIFICATION (DOTTED LINES)

In an investigation of the subjective effects of shock, Ss reported that some shocks from this device "pricked" the skin and were painful whereas other shocks of the same intensity and duration were not. An oscilloscope measurement showed that a sharp voltage spike occurred with the offset of each shock reported as painful but did not occur, or occurred at a lesser intensity, on shocks reported as not painful. With a shock level of 42 v. (*rms*) this spike was found to vary in amplitude between 0 and 127 v. (*max.*). Changing the position of Sw to the secondary transformer winding did not improve this condition.

This spike is related to the back *emf.* induced in the secondary transformer

* Received in the Editorial Office on January 18, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

winding when Sw is opened. The voltage amplitude of the spike depends upon the point in the AC cycle at which Sw is opened.

Since the back *emf.* varies in amplitude from shock to shock, the sensation produced will vary from shock to shock. For this reason the shock apparatus must be modified if it is to be satisfactory for use in most experiments.

Such modification is simple. A .0033 μfd paper capacitor (C in Figure 1), placed across the transformer secondary, bypasses the spike voltage without seriously affecting the 60 cycle AC voltage level.

*Department of Psychology & Psychiatry
University of Minnesota Medical School
Minneapolis 14, Minnesota*

EPIDEMIOLOGY OF MENTAL HEALTH IN COLLEGE*

Department of University Health, Yale University

RALPH M. RUST¹

A. INTRODUCTION AND PROBLEM

To me, the word epidemiology calls to mind a mental picture of a bare room with a desk and a detailed map on the wall behind it. At the desk sits a man deep in thought—perhaps haggard. Someone comes in with a report: "One more case of the plague has been diagnosed in South Haven." The haggard man gets up and puts another pin in the map. In my more imaginative moments, I might think of him quickly recalculating a statistic, one that would enable him to shout: "Aha! Now we have proof that the plague is caused by 'parentitis'!"

Such epidemiological fantasies are of course just that—fantasies. Even in specific contagious diseases, factors relevant or perhaps even pre-disposing to illness were soon recognized. Perhaps the epidemiological study of mental illness came directly from the recognition of the importance of such contingent variables—variables that make it more likely for one sort of person to be subject to the illness than another. Gordon, in the conference reported in *Epidemiology of Mental Disorder* (11), says, "epidemiology deals primarily with whole populations, the sick and the well, by establishing categories of persons and separating environmental factors to permit comparison and quantification of fractions of these populations with and without the particular disease."

Epidemiological studies of mental health (or mental illness) have become more frequent in the last decade or two and I will refer briefly to some of these later. However, few if any of these studies deal with a specific age or specific segment of the population. The epidemiology of mental health in college seems a particularly important area for study for a number of reasons. As Horst (8) suggests, useful information concerning an activity might be obtained from: "(a) those who have definitely achieved success in the activity, (b) those who have been unable to succeed in the activity, and (c)

* Received in the Editorial Office on January 25, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This paper is an expansion of part of a paper presented at a workshop on "Research in College Student Mental Health," Jules D. Holzberg, Chairman, American Orthopsychiatric Association Annual Meeting, 1959. The author is grateful to Robert L. Arnstein for his valuable editorial assistance.

persons who are learning the activity. . . . By studying people who are just beginning to engage in an activity, it is possible that more may be learned about the difficulties and abilities involved in it than by studying people who have had long experience in the activity." College students are all, in a sense, beginners, learning an adult rôle. Of course, they have really been doing this since childhood, but they are getting the chance to practice the adult rôle in a situation which is relatively free from parental demands and restraints. College students are in the process of "finalizing" some important decisions (marriage, career) and although it may seem that these decisions have already been made (as a function of the individual's class position, intelligence, etc.) there is still some room for modification and even where there is little room for change there's a certain stress that goes with actually deciding.

Besides being in the process of learning how to live as an adult, college students are more available for study than other groups. Often they will be spending four years at the same institution and can be easily followed. College students, because they are intelligent (and perhaps more interested in finding out about themselves) are better able to tell you what they think they're thinking and what they feel they're feeling. At Yale, for example, usually 75-80 per cent return the questionnaires which we have sent to undergraduates.

Perhaps what we learn from epidemiological studies of college populations will eventually enable us to identify certain characteristics of the college environment itself which interfere with the proper functioning of the student. It is possible too that such studies will help us select those students most likely to profit from the experience to be obtained at a particular college.

Defining mental health or the lack of it and determining the method by which "cases" are to be identified are two problems which immediately confront one in any epidemiological study of mental health.

I suppose one of my fantasies about the haggard man and the map is the idea that a case is easily defined—and another fantasy is that once defined a case is clearly identifiable. Certainly for mental health these are often confusing and difficult tasks. Identifying a case² involves coming to some kind of agreement about what a case is and finding some satisfactory way to assess individuals and to classify them as a case or a non-case. Perhaps in a certain sense everyone might be considered a case. Cole (2) states that, "All men are anxious, even those who love God and have been psychoanalyzed." The

² One section of papers presented at the 1952 annual conference of the Milbank Memorial Fund was devoted to this topic (12).

point at which one separates the case from the non-case may, in a certain sense, be relatively unimportant. This is particularly true if one accepts the concept of a mental health—mental illness continuum.

Farnsworth (+), reporting on recommendations of a conference of the World Federation for Mental Health lists nine things which mental health is not.³ "Mental health is *not*: (a) Adjustment under all circumstances. (b) Freedom from anxiety and tension. (c) Freedom from dissatisfaction. (d) Conformity. (e) Constant happiness. (f) A lessening of accomplishment and creativity. (g) The absence of personal idiosyncracies. (h) The undermining of authority. (i) Opposition to religious values."

In a more positive vein, Marie Jahoda (9) identifies what she perceives as mental health as: (a) correct perception of reality, (b) active adjustment (the ability to do something about problems with which one is faced), and (c) growth or unity of personality. The last named, perhaps because of the difficulty in defining it, has been dropped from later descriptions.

One problem that has remained unanswered is whether or not mental health is on a continuum with mental illness. There is a certain tendency on the part of all of us to speak of "mental health" or the "mentally ill." We tend to use similar terms for physical health, saying a person is in good health or poor health. Actually the degree of mental illness identifiable may be the result of several variables. There may be an underlying resistance or susceptibility to all illness. The specific illness (be it mental or physical) might be due to other factors. In some instances the presence of one disease might tend to preclude the presence of other diseases,⁴ while in other instances having one disease might make one more susceptible to some other. The total observable effect, however, suggests a continuum.

Clausen (1) points out that "we must ultimately face the task of relating mental health defined in positive terms to the individual's ability to resist mental illness under stresses which tend to precipitate such illness. Until this step has been taken, we shall not know whether positive mental health and mental illness are on a continuum or whether they are dimensions which crosscut each other as would many other aspects of personality. . . ."

At present, it seems generally more productive to study mental health by selecting cases in which mental health is absent rather than selecting cases

³ Joseph Wheelwright, M.D., Discussant for the workshop, kindly called my attention to these points.

⁴ Gregg (6) says that "angina pectoris and coronary sclerosis were given as a cause of death 14 times more often in the community than in state hospitals, and gastric and duodenal ulcer three times more often." In a subsequent study he concludes: "The incidence of disabling arthritis is 17 (times) greater in the community than among psychotics in Massachusetts state hospitals."

where it is present. By identifying those cases where some evidence for ill health exists, we seem more likely to arrive at a consensus than by studying cases that might be somewhat more dubiously selected as having good mental health.

Once some criteria have been established as to what constitutes a case there comes the problem of how to identify a case using the criteria established. The procedure used will depend in part on the availability and coöperativeness of the subjects.

Perhaps one of the easiest ways of identifying the case is some method which allows the case to identify himself. Hollingshead and Redlich (7) for example, selected as their sample for study those individuals from a certain area who were in psychiatric treatment on a certain date. This is a form of self-selection. The patients presumably came for treatment or were brought there by others. The study is well done and a valuable contribution, not only to the central problem (social class and mental illness) but to related problems. Like most studies however, it does not attempt to take into account those individuals in the community who are perhaps equally sick, but who are not in treatment. Studies which try to take into account both those in treatment and those not in treatment (total incidence or total prevalence studies) like the study of Roth and Luton (13) in Williamson County, Tennessee, or the Stirling County project under the direction of Leighton (10) are not easy to do. One of the primary difficulties is in evolving screening tests or interviewing techniques that will correctly assess individuals in the general population. With the house-to-house methods used in the Tennessee study almost twice as many cases were identified as had been previously identified in three years of intensive study in the community without such canvassing.

B. METHOD AND RESULTS

I would like to present one way of selecting cases and non-cases from among college students. The method is one which allows the subject to identify himself, but permits the identification of some subjects who, although they have disturbing personal problems, for one reason or another do not seek help with those problems. After describing the method of selecting cases and showing their distribution within one college (Yale), I will go on to show how cases differ from non-cases in three different kinds of data—in background variables, in the incidence of behavior they report, and in the adjectives they use to describe themselves.

Cases were selected on the basis of answers to two questions which were

part of a large questionnaire which the author and James S. Davie sent to a representative sample of all four undergraduate classes at Yale University. Students were asked: (a) how frequently they had been bothered by certain specific problems (loneliness, nervousness, insomnia, headaches, and indigestion), and (b) the extent to which any personal problem interfered with their activities (studies, athletic participation, extra-curricular participation, recreation, sleep, sex life, and relations with people). Using the twelve items from these two questions (the seven areas of functioning interfered with, and the five specific problems), each subject was scored one point for each item checked as bothering him very often or fairly often and one point for each item checked as seriously interfering or causing him some difficulty. These "Twelve-Problems" scale scores were used as a measure of lack of mental health. The responses to the items in the scale are discussed below

TABLE 1
DURING THE CURRENT SCHOOL YEAR, HOW OFTEN HAVE YOU BEEN BOTHERED BY:

Item	Per cent saying "Very often" or "Fairly often"
Loneliness	24
Nervousness	36
Insomnia	11
Headaches	9
Indigestion	8

TABLE 2
DO YOU AT PRESENT HAVE ANY PERSONAL PROBLEM WHICH INTERFERES WITH:

Area	Per cent replying "Yes, seriously interferes" or "Yes, some difficulty"
Studies	28
Athletic participation	19
Extra-curricular participation	19
Recreation	12
Sleep	14
Sex life	22
Relations with people	20

and are shown in Tables 1 and 2. In addition some information is presented on the relationship of the items in the scale to each other and to the total score on the Twelve-Problems scale. The data is based on the responses of the 833 subjects (65 per cent of those asked) who returned the questionnaire. The most frequently checked problem on the Twelve-Problems scale was "nervousness." Over one in three Yale students (36 per cent) say they are bothered by nervousness "very often" or "fairly often." It should be pointed

out that the subject was allowed to interpret the specific problem, and the response categories as he wished. The next most frequent problem (of those listed) was "loneliness," checked by almost one-quarter of the sample. The other three problems were less frequently checked: "insomnia" (11 per cent), "headaches" (9 per cent), and "indigestion" (8 per cent).

On the second question, designed to determine the specific areas of functioning interfered with rather than the frequency or nature of the problem, the area of functioning most frequently interfered with was "studies." Twenty-eight per cent of the sample indicated that they were currently bothered by a personal problem which seriously interfered with studying or gave them some difficulty. Approximately one-fifth of the subjects indicated that they currently had some personal problem which interfered with each of the following: "sex life," "relations with people," "athletic participation," and "extra-curricular participation." The other two areas were somewhat less frequently interfered with. "Sleep" was checked by 14 per cent and "recreation" was checked by 12 per cent of the group. It is not too surprising to find that the area which is most likely to be interfered with is the universal task of students, studying. Other areas such as sleep, recreation, and sex life (though they may perhaps be as universal) are apparently not so often considered to be interfered with.

In order to determine the extent to which these problem items measure the same or different variables, each of the 12 items was correlated with every other item and with three additional items.⁵ These additional items were: (a) "During the current school year have you ever consulted anyone about a personal problem?" (31 per cent answered "Yes"). (b) Was the student at any time a patient in the Student Mental Hygiene Clinic? (This information, obtained from official records, showed 12 per cent as patients at some time). (c) "Compared to other people your own age, how would you describe your physical health?" (Six per cent said, "A little worse" or "much worse").

Rather than present all these intercorrelations, I will attempt to summarize the important points. With few exceptions, the variables are all positively related. Correlations of areas interfered with range from .40 (athletic participation and sleep) to .82 (extra-curricular participation and recreation). The median correlation for this group of seven items was .57. Correlations for the group of items dealing with specific problems range from .24 (loneliness and insomnia, and loneliness and indigestion) to .71

⁵ For this analysis, estimated tetrachoric correlations were used. The complete matrix is available from the author on request.

(headaches and indigestion). The median correlation for this group of problems was .46. Correlations of items across the two questions ranged from $-.11$ (recreation and headaches) to $.75$ (insomnia and sleep), median correlation, $.27$. In general, across the two sets of items, "loneliness" seems to have the highest relationship to items in the areas of interference group; while "sleep" is one item in the areas of interference group most highly correlated with the specific problems group.

Turning from the relationship of the problems the subject has and the areas of functioning interfered with, to the area of seeking help, the two items with the highest relationship to being seen at the Student Mental Hygiene Clinic were having a personal problem interfering with studies (.47) and having a personal problem interfering with relations with people (.49). Being seen at the Clinic is correlated lowest with indigestion (.26) and with having a problem which interferes with athletic participation (.21). Consulting *anyone* about a personal problem during the current year correlates highest with loneliness (.34) and with having a problem which interferes with studies (.36), and correlates lowest with having a problem which interferes with one's sex life (.12) or sleep (.13) or being bothered by nervousness (.14) or insomnia (.13). As one would expect, those who tend to judge themselves as physically unhealthy compared to others their own age are most likely to be bothered by indigestion (.57) or headaches (.43). They are least likely to have a problem which interferes with their extra-curricular participation ($-.04$) or recreation (.00).

Table 3 shows the distribution of scores and the per cent of the total group scoring at each point on the Twelve-Problems scale. The range of scores for the 833 subjects was zero to 11. Almost three-quarters of the group indicated that they have had one or more bothersome personal problems during the

TABLE 3
THE TWELVE-PROBLEMS SCALE

Score	Number	Per cent
0	241	30
1	170	21
2	115	14
3	77	9
4	61	7
5	61	7
6	31	4
7	26	3
8	20	2
9	8	1
10	3	—
11	2	—

current year. About one-quarter of the group indicated that they have had four or more problems. An item analysis of the items in the scale was performed by correlating each item against the total score on the scale. All but two individual items were correlated above .70 with total score. These two items, nervousness and headaches, correlated .62 and .68 with total score. Having a personal problem interfering with relations with people had the highest correlation with total score (.96).

Assuming that the scores on the Twelve-Problems scale provide some measure of the mental health—mental unhealth continuum, the author was interested in assessing some of the variables that might be related. Three general classes of data were used for this assessment. First there are certain "background variables" that were collected on all subjects. These variables are in part self-report items and in part based on university records. These variables are shown in Table 4. Second, data were drawn from Reported Behavior Inventory scales. These scales, 17 in all, were designed to measure various aspects of student behavior and were developed⁶ from 220 items ranging from innocuous items such as "Have you ever given blood?" and "Have you ever been inside the Yale Library?" to very personal items such as "Have you ever engaged in a homosexual act?" and "Have you ever had sexual intercourse?" A third comparison is based on responses to a 160-item adjective check-list, where subjects were asked to check each adjective they considered more true of them than untrue and cross out those they considered more untrue than true.

For purposes of this analysis the group was divided into approximately equal fourths, according to the number of problems indicated on the Twelve-Problems scale: (a) zero problems (241 subjects), (b) one problem (170 subjects), (c) two or three problems (192 subjects), and (d) four or more problems (212 subjects). The results are shown in Tables 4, 5, and 6.

Table 4 shows the background variables that are significantly related to the number of problems as measured by the Twelve-Problems scale. Only those variables which show significant differences (using chi square, $p < .01$) among groups and which were judged to be "reasonably" linear are included in the table. The results show that the greatest differences are in terms of satisfaction the individual feels with his experience at Yale.⁷ Those with the most problems are the least likely to feel satisfied with their experience. Also they regard themselves as less healthy physically and are more likely to have been seen in the Student Mental Hygiene Clinic. Those with more problems

⁶ By the author and James S. Davie.

⁷ A similar finding for Clinic users (as opposed to non-users) was found by James S. Davie (3).

are apparently somewhat more likely to have higher Scholastic Aptitude Test scores (Verbal section), but are less likely to report good study habits (14). Those with *no* problems are more likely to spend above the median amount of time in extra-curricular activities. Differences on other variables tested (these are listed at the bottom of Table 4) are small and/or unclear as to direction.

The results of relating the Reported Behavior Inventory scales to the

TABLE 4
BACKGROUND VARIABLES AND THE TWELVE-PROBLEMS SCALE

Variable	Number of problems			
	0	1	2-3	4 or more
	Per cent			
Consulted Student Mental Hygiene Clinic	05	10	11	25
Extra-curricular activities (11 or more hours per week)	51	38	41	38
Satisfaction with current academic situation (very or fairly satisfied)*	71	58	52	45
Satisfaction with current non-academic situation (very or fairly satisfied)*	75	61	61	45
High Satisfaction Scale score (Davie)	73	59	47	29
High Scholastic Aptitude Test score (600 or higher)	44	51	61	56
Good Study Habits (Rust)	54	46	36	27
Self-rating of physical health (a little worse or much worse than average)	03	04	07	13

* These two items summed also showed significant differences among the four groups as did cumulative academic and non-academic satisfaction.

Note: Other variables tested and showing non-significant or essentially non-linear differences were: Class, school and degree, type of secondary school, estimated social class and social mobility, Yale relatives, financial aid status, estimated study time, predicted average, average grade, achievement status.

TABLE 5
REPORTED BEHAVIOR INVENTORY SCALES AND THE TWELVE-PROBLEMS SCALE

Reported behavior inventory scale*	Number of problems			
	0	1	2-3	4 or more
	Per cent with high score			
Health	65	63	49	41
Involvement with Yale	62	54	58	42
Athletic Activity	54	50	49	34
Athletic Spectator Interest	69	64	65	47
Male Peer Relationships	56	51	45	32
Relations with parents	64	62	53	45
Social Relations with girls	57	50	45	42

* All scales were developed by the author and James S. Davie. Scores are for current behavior. A high score is positive, e.g., better health, more athletic activity, etc.

Note: Other scales showed non-significant or essentially non-linear differences. These were: Drinking, Contact with the Faculty, Self-Propelled Intellectual Activity, Study Habits, Academic Standing, Spectator Interest in the Arts, Non-athletic Active Recreation, Heterosexual Activity, Moral Behavior, and Work Experience.

Twelve-Problems scale are shown on Table 5. The criteria for including these particular scales as significant are the same as for the background variables. Seven of the 17 scales yield significant differences among groups. In all seven instances, those with fewer problems are more likely to score above the median on the Reported Behavior Inventory scale under consideration. It should be noted that the point of greatest difference is usually

TABLE 6
THE ADJECTIVE CHECK-LIST AND THE TWELVE-PROBLEMS SCALE

Adjectives characteristic of those with problems					Adjectives characteristic of those without problems				
Adjective	Number of problems				Adjective	Number of problems			
	0	1	2-3	4+		0	1	2-3	4+
Per cent checking					Per cent checking				
Lonely	08	20	33	62	Contented	81	70	52	35
Worrying	24	46	55	74	Well-balanced	95	88	81	54
Insecure	10	17	33	58	Satisfied	69	60	49	30
Tense	19	33	38	62	Happy	96	87	82	58
Moody	36	53	55	78	Relaxed	80	69	56	42
Dissatisfied	24	34	49	65	Calm	81	63	60	49
Highstrung	21	24	32	54	Integrated	73	69	65	41
Immature	15	19	22	48	Consistent	74	67	62	46
Restless	47	60	65	77	Conventional	62	61	45	34
Withdrawn	13	18	23	42	Efficient	76	69	55	48
Anxious	55	55	66	83	Mature	81	81	76	54
Nervous	23	42	51	62	Confident	82	73	68	57
Changeable	57	66	69	84	Poised	76	75	63	53
Inhibited	30	34	44	57	Optimistic	81	71	69	61
Easily misunderstood	23	23	31	49					
Rebellious	25	28	29	51					
Cynical	42	45	52	67					
Day-dreaming	57	70	70	80					
Irritable	30	34	40	53					
Shy	27	31	37	50					
Guarded	29	28	36	51					
Awkward	07	10	17	28					

between those with two or three problems and those with four or more. This may very likely be so because the "distance" between the group with four or more problems (with an average of 5.67 problems) and the next lower group (with an average of 2.40 problems) is greater than that between any other adjacent groups. It is not surprising that those with fewer problems score higher on the Health scale (have better health) since this scale includes items dealing with use of the Health Department and the Clinic. Other differences are less obvious, but not too surprising. Those with more problems seem to have poorer relations with parents, less social experience with

girls (although they do not differ in heterosexual experience), and are less successful in relationships with male peers. There is less involvement with Yale, less spectator interest in sports, and less athletic activity. Other scales (listed at the bottom of Table 5) show differences among groups which are smaller or essentially non-linear. It should be pointed out that there are apparently no real differences among these groups in grades or in academic standing, even though studying is presumably one activity frequently interfered with.

The adjective check-list showed greater differences between those scoring high and those scoring low on the Twelve-Problems scale than either the background variables or the Reported Behavior Inventory scales. Perhaps this is because both the Twelve-Problems scale and the adjective check-list are self-evaluative instruments, less objective than either of the other kinds of data.

Of the 160 adjectives on the check-list, two-thirds (106) were significantly related to scores on the Twelve-Problems scale ($p < .05$). Those 36 adjectives with a difference of 20 per cent or more between the proportion of those with zero problems and the proportion of those with four or more problems checking a given adjective are shown in Table 5. These differences are all clearly significant ($p < .01$). The range of words characteristically checked by those with problems gives an interesting word picture of the subjects in this group. They describe themselves as lonely, worrying, insecure, moody, and tense. Those without problems are more likely to describe themselves in ways which are consistent with lack of problems—contented, well-balanced, calm, happy, integrated, relaxed, and satisfied.

C. DISCUSSION

The method of case selection used in this study (the Twelve-Problems scale) seemingly provides a way of distinguishing cases and non-cases which is quite useful as a research tool. This method, based on the self-identification of cases by means of a questionnaire, seems particularly useful because it selects cases that would not ordinarily come to the attention of clinics and hence does not confound willingness to come to a clinic with having problems (as do those methods which rely on coming to a clinic as a method of case selection). The method still does not include for study those individuals who have problems but who are either unaware of these problems or unwilling to admit them.

For descriptive purposes the method seems quite useful. The data on background variables, Reported Behavior Inventory scales, and on the

adjective check-list yielded differences among the groups studied that seem internally consistent and in line with general expectations. For descriptive purposes then, when the instrument is used as a research tool on college students and not for screening, the Twelve-Problems scale seems useful and effective. Whether it is as useful for descriptive purposes as other methods, such as a patient/non-patient criterion or as some other methods of self-selection, is a problem for further research. So, too, are the usefulness of the Twelve-Problems scale for groups other than college students and as a screening device.

Although the reported results seem to provide a good description of groups differing in terms of the mental health continuum, very little information was provided about the etiology of personal problems. Perhaps there were actually too few variables studied that could be related to etiology, perhaps the failure of the present method to select those subjects who have problems but do not report them, and perhaps the inadequacies of a cross-sectional as compared with a longitudinal approach all contribute to the difficulty.

As in any research effort the present study suggests more questions than it answers. The fact that only 25 per cent of the students with four or more problems on the Twelve-Problems scale were seen in the Student Mental Hygiene Clinic raises a number of practical and theoretical questions. How do those who came to the Clinic differ from those who did not come? Are those who did not come equally disturbed by their problems? Do they take their problems somewhere else? Or is a differential willingness to seek assistance the crux of the matter?

Another interesting question for further study might be an investigation of those former patients at the Mental Hygiene Clinic who report four or more problems. In a certain sense they might be regarded as therapeutic failures. Perhaps a more clinical study of this group might yield some interesting and useful information. It is encouraging to note that only one-third of the "former" patients (students who were seen in years prior to the year the questionnaire was given) as compared with two-thirds of the "current" patients (students seen during the year the questionnaire was given) have four or more problems. This difference may or may not be due to psychotherapy, but it is plausible to assume that psychotherapy has helped.

Some tentative approaches to the study of the effects of psychotherapy are also suggested. This is a difficult area and one in which motivation for treatment has to be carefully considered as a variable. It might be enlightening however, to select a group of freshmen from the present study none of whom were seen at the Clinic during their freshman year and compare those

who were subsequently seen in the Clinic with another group of subjects from the same class who reported an equal number of problems but who did not subsequently come to the Clinic. At some later date, they could be retested to determine if there were differences in the number of problems they then reported.

These and many other studies are still to be done, although most of them are not easy or inexpensive. Nevertheless, college mental hygiene clinics must continue to do research in this area. Hopefully, such research will lead to a better understanding of the nature, severity, prevalence, and perhaps even the etiology of problems of college students and what can be done about them.

D. SUMMARY

Because college students are in the process of learning to live as adults and are "finalizing" some important life decisions, and because they are coöperative and intelligent, they seemed a particularly appropriate choice for the epidemiological study of mental health. The variable of mental health or the lack of it was assumed, at least for practical purposes, to lie on a continuum and was evaluated by means of a "Twelve-Problems" scale developed by the author. Data for the study was obtained from questionnaires administered to 833 Yale University undergraduates.

The paper describes the development of the Twelve-Problems scale and the intercorrelation of the 12 items in the scale. Thirty per cent of the group studied indicated no problems and one-quarter reported four or more problems.

College students with relatively poor mental health (as measured by the Twelve-Problems scale) were compared with those with relatively good mental health on three types of data: background variables, Reported Behavior Inventory scales, and a self-descriptive adjective check-list. On the background variables, those students with relatively poor mental health are less satisfied with their college experience, have poorer study habits, rate themselves as less healthy physically, and, as might be expected, a higher proportion of them have visited the Mental Hygiene Clinic (although three-quarters of those with four or more problems have not been seen at the Clinic). On the Reported Behavior Inventory scales, those students with poorer mental health have poorer interpersonal relationships (with parents and with male and female peers), are less involved with their college, and are less interested in things athletic. On the adjective check-list students with problems more frequently check those adjectives which are indicative of a self-critical attitude as well as those which are consistent with the presence of psychological problems.

It was suggested that the Twelve-Problems scale might be useful for epidemiological studies of other groups as well as for more intensive studies of college students.

REFERENCES

1. CLAUSEN, J. A. *Sociology and the field of mental health*. New York: Russell Sage Foundation, 1956.
2. COLE, W. G. *Sex in Christianity and psychoanalysis*. New York: Oxford Univ. Press, 1956.
3. DAVIE, J. S. Who uses a college mental hygiene clinic. In B. M. Wedge (Ed.), *Psychosocial Problems of College Men*. New Haven: Yale Univ. Press, 1958. (Pp. 140-149.)
4. FARNSWORTH, D. L. *Mental health in college and university*. Cambridge, Mass.: Harvard Univ. Press, 1957.
5. FUNKENSTEIN, D. H., & WILKIE, G. H. *Student mental health: An annotated bibliography, 1936-1955*. London: World Federation for Mental Health.
6. GREGG, D. The paucity of arthritis among psychotic cases. *Amer. J. Psychiat.*, 1939, **95**, 853-861.
7. HOLLINGSHEAD, A. B., & REDLICH, F. C. *Social class and mental illness*. New York: Wiley, 1958.
8. HORST, P. The prediction of personal adjustment. *Soc. Sci. Res. Coun. Bull.*, 1941, No. 48.
9. JAHODA, M. Toward a social psychology of mental health. In M. M. E. Senn (Ed.), *Problems of Infancy and Childhood*. New York: Josiah Macy, Jr. Foundation, 1950.
10. LEIGHTON, A. H. Psychiatric disorder and social environment. *Psychiatry*, 1955, **18**, 367-383.
11. MILBANK MEMORIAL FUND. *Epidemiology of mental disorder*. New York: Milbank Memorial Fund, 1950.
12. ———. *Interrelations between the social environment and psychiatric disorders*. New York: Milbank Memorial Fund, 1953.
13. ROTH, W. F., & LUTON, F. H. The mental health program in Tennessee: Statistical report of a psychiatric survey in a rural county. *Amer. J. Psychiat.*, 1943, **99**, 662-675.
14. RUST, R. M. Personality and academic achievement: A questionnaire approach. In B. M. Wedge (Ed.), *Psychosocial Problems of College Men*. New Haven: Yale Univ. Press, 1958. (Pp. 45-62.)

Department of Health
Yale University
435 College Street
New Haven, Connecticut





TIME PERCEPTIONS AS AFFECTED BY NEED TENSION*

Department of Psychology, Trinity College, Hartford, Connecticut

ROBERT D. MEADE¹

A. PROBLEM

Rosenzweig and Koht (4), in a widely-quoted study, reported that time estimates are shorter when a condition of "need tension" is present in *S* than when such need is absent. Conditions of high and low need tension in this study were brought about in essentially the same way as what Iverson and Reuder (2) call the intelligence test deception method of ego-involvement. However, most of the *Ss* were tested first under low need and then under high need leaving open the possibility of confounding of results by sequential effects such as familiarity. In addition, no statistics are reported concerning the reliability of the findings.

Recent work by Hindle (1) and Meade (3) suggests that a factor other than that of need tension could explain the Rosenzweig and Koht findings. Hindle has found that time estimates are inversely related to perceived rate of progress through a task while Meade has shown that this effect holds only where there is motivation to reach the end of the task.

It is believed that *Ss* who work on two similar tasks or puzzles in succession, as was the case in the Rosenzweig and Koht study, will be more familiar with such puzzles on the second than on the first and this would make *Ss* more capable of perceiving progress on the second than on the first. This would, then, be a condition for time estimates to be shorter on Puzzle 2. Having the low need condition come first followed by the high need condition, as in the earlier study, would confound the effects of motivation with that of perceived progress. It is predicted that control or elimination of the familiarity factor would modify or eliminate the trend reported by Rosenzweig and Koht. Two experiments were done to test this prediction. One study repeated the earlier experiment but counterbalanced the order of experimental conditions to cancel out sequential effects. Such a procedure, however, presents a possible difficulty—the question of the validity of creating a condition of low need tension following high need tension which was neces-

* Received in the Editorial Office on January 27, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This investigation was supported by a research grant, M-3193, from the National Institute of Mental Health, Public Health Service.

sary for half of the *Ss* in this experiment. Consequently, a second study using two random groups each under a different experimental condition was done.

B. METHOD

1. *Experiment 1*

a. Subjects. Sixty college students (30 men and 30 women).

b. Apparatus. Two unsolvable wooden block puzzles similar to those described in the Rosenzweig and Koht study.

c. Procedure. All *Ss* were tested individually and were not aware, until the end of the experimental session, that they were participating in a time perception study. Before the experiment began, *E* asked *S* to lend him his watch, if he had one, on the pretext that his own was broken and it was needed as part of the procedure. *Ss* were arbitrarily separated into two sub-groups, 15 men and 15 women in each—one sub-group being tested first in the low need tension condition and then in the high, the other group being tested in the reverse order. *Ss* in the first sub-group were given the following instructions:

In a little while I am going to give you an intelligence test. It will consist of some puzzles, like this one, which you are to solve. For example, in the puzzle in front of you now, the wooden blocks when fitted together properly make a perfect square eight inches on a side. Before we can start, however, I need to bring some materials from another room and this will take a little while. I would like for you to work on this sample puzzle for practice while I am gone.

E left the experimental room and returned exactly 5 min. later and said, "How long do you estimate that I was gone?" After obtaining *S's* time estimate, the instructions continued.

I am now ready to start the intelligence test itself. I need to know how bright you are so that I can determine whether you are intellectually capable of doing some other problems I will have in a few weeks. In addition, you, yourself, may be interested in your intelligence scores. As I told you a few minutes ago, the parts of the test consist of solving puzzles like the one you have just tried. The first puzzle of the test will also make a perfect square 8 in. on a side. Here it is. It is important for you to work as rapidly as possible, since all parts of the tests are timed.

After answering any questions, *E* produced a stopwatch and started *S* working on the puzzle. When *S* had been at work for exactly 5 min., *E* stopped him and asked, "How long do you estimate you were working on this puzzle?"

The second sub-group of *Ss* were given the high need instructions first and

a time estimate was obtained for a 5-min. period. Next, *E* told *S* that there would be a rest period and that the test would be resumed shortly. First, however, *E* said that it was necessary to obtain some additional materials from another room. *S* was told that in the meantime he could practice on another puzzle until *E* returned. This practice puzzle was placed in front of *S* and *E* left the room. *E* returned exactly 5 min. later and obtained the time estimate for the period of his absence as in the first sub-group.

2. Experiment 2

a. Subjects. Two groups of 30 randomly-selected male college students.

b. Apparatus. Two wooden block puzzles described in Experiment 1.

c. Procedure. All *Ss* were tested individually and were naïve concerning the nature of the experiment. Their watches were taken away during the conduct of the experiment as in Experiment 1. Group 1 was tested under the low need tension only while Group 2 was tested under the high need tension condition only. Instructions for the high and low need conditions were the same as those used in Experiment 1 and all time periods were 5 min.

C. RESULTS

As can be seen from Table 1, there is no significant difference in time estimates between the high and low need tension conditions in either of the experiments and this supports the predictions made in the introduction.

TABLE 1
EFFECT OF HIGH AND LOW NEED TENSION ON TIME ESTIMATES

Experiment	Objective time	Mean high need	Mean low need	<i>t</i>	<i>p</i>
1	5 min.	5.4	5.1	.46	> .05
2	5 min.	5.3	5.5	.31	> .05

Statistical evaluation of the effect of order of presentation of the two conditions of the independent variable in Experiment 1 resulted in a mean estimate of 5.6 min. for the condition which appeared first and 4.9 min. for the second condition. This difference resulted in a *t* of 1.71 which is significant at the .05 level of confidence (one-tailed test). Rosenzweig and Koht reported that there was a tendency for female subjects to estimate the high need period as shorter more than male subjects. This effect was not found in the present investigation.

D. DISCUSSION

The results of this experiment indicate that need tension alone does not affect time perceptions as was reported by Rosenzweig and Koht. The significantly longer estimates for the second of the two periods in Experiment 1 supports the view that progress rather than motivation was responsible for the earlier results.

In both Experiment 1 and the Rosenzweig and Koht study no control of perceived progress was effected and any information about progress *S* received from his own work was probably ambiguous. However, it is believed that experience with the first puzzle could create enough familiarity to allow *S* to determine for himself, to some extent, when he was making progress. Thus, it would be expected that time estimates would be shorter in the second puzzle.

Meade (3) has found that perceived rate of progress affects time estimates when there is motivation to reach a goal but not when motivation is absent or at a very low level. It can be argued that in the present study as well as the earlier one there was motivation, of a considerable degree, in both the high and low need conditions. *Ss* under both conditions were told that the puzzles on which they worked would be used to measure intelligence. Further, it would seem to be difficult to tell, in this situation, whether there is any difference in the degree of motivation between the two conditions. It is not known whether *Ss* have a greater motivation while performing a task for practice than while doing the same kind of task as a part of an intelligence test. It is also possible that other kinds of motivation entered the experimental situation such as motivations to please or to follow *E's* directions.

From these considerations, further support for the contention that need tension, as employed here, does not affect time estimates is adduced. It is predicted that in a factorial experiment where high and low need tension and perceived rate of progress are used as independent variables, time estimates will be inversely related to rate of progress but will be independent of need tension. Tests of this prediction are now being conducted at this laboratory.

E. SUMMARY

Rosenzweig and Koht (4) reported that time estimates are shorter under high need tension than under low need tension. Two modified experiments were done to test this conclusion. In Experiment 1, 60 *Ss* worked on wooden block puzzles for 5 min. under high and 5 min. under low need tension with the order of conditions counterbalanced. In Experiment 2, one group of 30 *Ss* worked on one puzzle 5 min. under high need tension while the other group of 30 *Ss* worked 5 min. under low need tension.

There was no significant difference between high and low need tension conditions with respect to time estimates in either experiment. In Experiment 1, it was found that the first condition was estimated significantly longer than the second. This supports the contention that perceived progress rather than need tension was responsible for the earlier results.

REFERENCES

1. HINDLE, H. Time estimates as a function of distance traveled and relative clarity of a goal. *J. Personal.*, 1951, **19**, 483-501.
2. IVERSON, M., & REUDER, M. Ego involvement as an experimental variable. *Psychol. Rep.*, 1956, **2**, 147-181.
3. MEADE, R. Time estimates as affected by motivational level, goal distance and rate of progress. *J. Exp. Psychol.*, 1959, **58**, 275-279.
4. ROSENZWEIG, S., & KOHT, A. The experience of duration as affected by need tension. *J. Exp. Psychol.*, 1933, **16**, 745-775.

Department of Psychology
Trinity College
Hartford 6, Connecticut

TEST ITEM DIFFICULTY AND THE RELIABILITY OF ITEM ANALYSIS METHODS*

Department of Psychology, Temple University

JAMES F. ADAMS¹

A. PROBLEM

Considerable attention has been focused upon the optimum level of test item difficulty for maximizing *test* reliability and validity. Less attention has been directed towards the reliabilities of the *methods* of item analysis used in the selection of test items. This study deals with the latter problem, and in particular, the reliability of item analysis methods as a function of test item difficulty. The following methods of item analysis are examined: the U-L 27 per cent method (7, 8), the U-L 10 per cent method (2), phi (11), tetrachoric r (4), tetrachoric r corrected for non-median dichotomization (14), biserial r , point biserial r , the t -ratio (19), and the discrimination index (5). Additional problems considered are the effect of item difficulty level on the validity of test items and the interrelationships between item analysis methods.

B. METHOD

1. *Test Items*

The final examinations for a two-semester course in introductory psychology were obtained for purposes of item analysis. The examination for each semester had 140 four-alternative multiple choice items. The test items were balanced evenly over the first eight chapters of the text (16) for the first semester and the last eight chapters of the text for the second semester.

The reliability of the examinations was computed by using the Kuder-Richardson formula No. 21. As this formula assumes that all of the items of a test have approximately the same degree of difficulty, the computed coefficients of .86 for the first semester examination and .83 for the second semester examination, probably underestimates the actual reliabilities of the two tests.

In scoring the examinations no penalty was employed for guessing. The

* Received in the Editorial Office on January 29, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This paper is based upon a dissertation submitted to the State College of Washington in partial fulfillment of the requirements for the degree of doctor of philosophy in psychology. The study was supported by a grant from The James McKeen Cattell Fund.

students were encouraged to thoughtfully answer every item, and every student answered every item in the two-hour examination period allotted for testing in each semester.

2. *Subjects*

The subjects used in this study were those students in the introductory psychology course at Miami University who took the final examination at the end of each semester. The first semester examination was taken by 720 students and the second semester examination by 640 students. The majority of the students were of freshman and sophomore standing.

3. *Procedure*

The procedure in this study was identical in the treatment of both examinations; i.e., a replication of procedure was used in the second half of the study.

After the examinations were scored, the papers were ordered by total examination score (the criterion). Two item analysis groups were constructed by taking every other examination paper from the ordered series. The groups for the first examination were designated *A* and *B*. The second examination item analysis groups were designated *C* and *D*. The respective mean test scores for Groups *A* and *B* were 99.90 and 99.98; the standard deviations were 15.07 and 15.06. The mean test scores for Groups *C* and *D* were 103.50 and 103.48. The standard deviations were 13.50 and 13.68.

The difficulty level of an item was determined by the percentage passing the item in the combined groups. Originally it had been planned to allocate the test items to five difficulty categories. As there were few very difficult items, four subtests were used with difficulty characteristics of 0-40 per cent passing, 41-60 per cent passing, 61-80 per cent passing, and 81-100 per cent passing. The reliability of item difficulty level was also computed and found to be quite substantial.

The mean item validity indices for the difficulty subtests and for the total test was computed for each group and for every method of item analysis used. It should be recognized that the term "item validity" as it is here used is, in effect, homogeneity. As the distribution of r is skewed, all item-criterion correlations were converted to Fisher's z . Further statistical computations involved the use of z rather than r .

The reliability of each item analysis method was determined by correlating the item indices obtained in Group *A* with the results obtained from using the method in Group *B*. Reliability was determined for each of the difficulty

subtests and for the total test of 140 items. The same procedure was followed with Groups *C* and *D* with the second examination.

The intercorrelations between item analysis methods was found by averaging the item indices for a given method on each item for both groups, e.g.,

$$\bar{X}_{AB} = \frac{A_i + B_i}{2}$$

for the first semester exam. These values for the 140-item test, separately computed for each method of item analysis, were then intercorrelated.

C. RESULTS

The mean values of the item-criterion correlations, *t*-ratios, and discrimination indices are listed by difficulty level in Table 1. As the indices were comparable for Groups *A* and *B* in the first semester and *C* and *D* in the second semester and showed the same trends in both semesters, i.e., the replication of procedures produced comparable results, the four groups have been combined in Table 1.

TABLE 1
COMBINED MEAN ITEM-CRITERION CORRELATIONS, *t*-RATIOS, AND DISCRIMINATION
INDICES AT VARYING DIFFICULTY LEVELS FOR GROUPS A + B + C + D

Diff. Level	Item N	U-L 27%	U-L 10%	Phi	Tetra.	Tet. Cor.	Biser.	Pt. Bis.	<i>t</i> -ratio	Discrim.
0-40%	44	.24	.24	.14	.24	.24	.24	.17	3.25	15.
41-60%	82	.36	.36	.23	.36	.36	.34	.27	5.12	22.
61-80%	202	.41	.41	.24	.44	.41	.40	.30	5.69	26.
81-100%	232	.40	.37	.18	.54	.44	.46	.24	4.65	26.
0-100%	560	.39	.38	.21	.46	.40	.40	.26	4.98	24.

Note: The combined subject *N* for Groups A + B + C + D = 1360. The Item *N* is the combined Item *N* for Groups A + B + C + D at the given difficulty level.

Method reliability, as used in this study, refers to the extent that an item analysis method produces comparable item-criterion correlations (*t*-ratios or discrimination indices) on the same test but with equated groups. The reliability of item analysis methods was computed for both semester examinations; however, inasmuch as the findings were quite similar for both semesters, the results have been combined for Table 2, i.e., the method reliability coefficients of Table 2 are the means of the coefficients from both semesters.

The significance of the differences between method reliability coefficients for the difficulty subtests and the total test (Table 2) are reported in Tables 3 and 4. It should be noted that the entries of Tables 3 and 4 are normal deviates (6, pp. 305-306) and were evaluated by reference to the table of the

TABLE 2
RELIABILITY OF ITEM ANALYSIS METHODS AS A FUNCTION OF TEST ITEM
DIFFICULTY—SEMESTERS COMBINED

Diff. Level	Item N	U-L 27%	U-L 10%	Phi	Tetra.	Tet. Cor.	Biser.	Pt. Bis.	t-ratio	Discrim.
0-40%	22	.66	.71	.65	.54	.58	.67	.70	.74	.66
41-60%	41	.84	.79	.84	.80	.83	.91	.91	.91	.83
61-80%	101	.78	.63	.77	.76	.75	.81	.81	.82	.77
81-100%	116	.61	.54	.51	.47	.45	.59	.60	.55	.60
0-100%	280	.74	.66	.72	.65	.66	.74	.78	.78	.72

normal curve. As in each case the two z 's evaluated are based on the same sample while the test of significance used assumes different samples, it is possible that the table entries *underestimate* the actual significance of the differences. That is to say, the entries in Tables 3 and 4 may be conservative estimates of the true differences between the method reliability coefficients.

The intercorrelations between the item analysis methods used on the examination for each semester are cited in Table 5. As the discrimination index is a transformation of the U-L 27 per cent method and was found to correlate perfectly with this method, it was not included in the table.

TABLE 3
SIGNIFICANCE OF DIFFERENCES BETWEEN RELIABILITY COEFFICIENTS FOR ITEM
ANALYSIS METHODS AT VARYING LEVELS OF ITEM DIFFICULTY

Methods	U-L 27%	U-L 10%	Phi	Tetra.	Tet. Cor.	Biser.	Pt. Bis.	t-ratio	Discrim.
U-L 27%		.83	1.13	1.50	1.73	.23	.15	.68	.15
U-L 10%	.65	1.94	.13	.26	.45	.58	.58	.77	.13
Phi	.31		.30	.68	.90	.60	.68	.15	.68
	.00	.65	1.81	1.68	1.48	2.52	2.52	2.71	1.81
Tetra.	.03	.34		.38	.60	.90	.98	.45	.98
	.52	.13		.13	.32	.71	.71	.90	.00
Tet. Cor.	.59	.89	.52		.23	1.28	1.35	.83	1.35
	.13	.52	.13		.19	.84	.84	1.03	.13
Biser.	.40	.71	.37	.39		1.50	1.58	1.05	1.58
	1.35	2.00	1.35	.18		1.03	1.03	1.23	.32
Pt. Bis.	.06	.25	.09	1.87	1.48		.08	.45	.08
	1.35	2.00	1.35	.65	.46		.00	.19	.71
t-ratio	.25	.06	.28	1.87	1.48	.00		.53	.00
	1.35	2.00	.83	.65	.18			.19	.71
Discrim.	.49	.18	1.35	1.87	1.48	.00	.00		.53
	.13	.52	.52	1.08	.89	.43	.25		.90
	.00	.30	.13	.39	.00	1.48	1.48	1.48	
			.03	.59	.40	.06	.25	.49	

Note: This table should be interpreted as follows. The lower half of the table contains the analyses for items of 0-40% difficulty (lower number in cell) and items of 41-60% difficulty (upper number of cell). The upper half of the table contains the analyses for items of 61-80% difficulty (lower number in cell) and items of 81-100% difficulty (upper number in cell). A z of 1.65 is needed for significance at the .10 level; 1.96 at the .05 level, and 2.58 for significance at the .01 level.

TABLE 4
SIGNIFICANCE OF DIFFERENCES BETWEEN RELIABILITY COEFFICIENTS FOR ITEM
ANALYSIS METHODS AT THE 0-100 PER-CENT LEVEL OF ITEM DIFFICULTY

	U-L 27%	U-L 10%	Phi	Tetra.	Tet. Cor.	Biser.	Pt. Bis.	t- ratio	Dis- crim.
U-L 27%		1.88	.47	2.00	1.88	.00	1.06	1.06	.47
U-L 10%			1.41	.12	.00	1.88	2.94	2.94	1.41
Phi				1.53	1.41	.47	1.53	1.53	.00
Tetra.					.12	2.00	3.06	3.06	1.53
Tet. Cor.						1.88	2.94	2.94	1.41
Biser.							1.06	1.06	.47
Pt. Bis.								.00	1.53
t-ratio									1.53
Discrim.									

Note: A z of 1.65 is needed for significance at the .10 level; 1.96 for significance at the .05 level; and, 2.58 for significance at the .01 level.

TABLE 5
INTERCORRELATIONS BETWEEN ITEM ANALYSIS METHODS

Methods	U-L 27%	U-L 10%	Phi	Tetra.	Tet. Cor.	Biser.	Pt. Bis.	t-ratio
U-L 27%		.87	.85	.75	.86	.76	.92	.89
U-L 10%	.92		.83	.58	.72	.68	.91	.89
Phi	.94	.86		.55	.76	.58	.94	.93
Tetra.	.82	.75	.68		.96	.84	.64	.62
Tet. Cor.	.90	.84	.84	.97		.85	.80	.79
Biser.	.88	.86	.76	.94	.95		.73	.78
Pt. Bis.	.95	.94	.95	.77	.88	.88		.97
t-ratio	.96	.94	.96	.77	.89	.88	1.00	

Note: The lower half of the table contains the data from the first semester examination. The upper half contains the data from the second semester examination.

D. DISCUSSION

In Table 1 it is of interest to note that the mean item-criterion correlations for the *U-L* 27 per cent method, the *U-L* 10 per cent method, tetrachoric r corrected for non-median dichotomization, and the biserial r are quite similar. While phi tends to run somewhat lower than the point biserial r , these indices are also comparable. The t -ratio, although a test of significance rather than a coefficient of correlation, follows the same pattern as does phi and the point biserial coefficient.

It would appear from Table 1 that there is a trend for the easier items to be the "best" items. Although Nefzger and Drasgow (13) have pointed out that Pearson's product moment coefficient does not assume normally distributed variables, item analysis techniques which are intended to give estimates of the product moment coefficient, or, *are* the product moment coefficient for a particular situation, need to be interpreted cautiously if the criterion scores

are not normally distributed. Adams (1) has commented upon the effect of non-normally distributed criterion scores on item analysis methods. Departures from normality can seriously affect the size of the item-criterion correlations *regardless* of the assumptions or lack of assumptions on this variable. This is particularly true as test items increase in validity. All of the criterion score distributions in the present study were skewed in the direction of the low scores. It is likely that this accounts in a large measure for the seeming trend towards "better" items at the easier levels although it is unlikely that it accounts for all of the reduced mean magnitudes of the most difficult items.

The effect of non-normal criterion scores upon methods of item analysis has not received adequate recognition in the literature. Cook (3), Smith (17), Long and Sandiford (12), Richardson (15), and Vernon (18) have all reported a correlation between difficulty level and the size of the biserial correlation coefficient although this method is theoretically independent of this factor. In every case where it is possible to examine the distributions of criterion scores for these studies it should be noted that the scores have varying degrees of skewness. This writer has concluded, after examining a number of subject matter examinations from a number of different disciplines, that skewness is more typical than normality. If this is true, skewness should be taken into consideration when evaluating test items at varying difficulty levels.

From the data with respect to the reliability of item analysis methods (Tables 3 and 4), the following conclusions can be made: The reliability of every method is the lowest for the easiest items. Items in the middle range of difficulty have the highest method reliability, followed closely by items in the 61-80 per cent category. There are no significant differences between method reliability coefficients for the most difficult items. At the 41-60 per cent difficulty level, biserial r , point biserial r , and the t -ratio have the highest method reliability. The same is true for the 61-80 per cent category. The methods which compare the least favorably are the $U-L$ 10 per cent method, tetrachoric r , and tetrachoric r corrected for non-median dichotomization. The $U-L$ 27 per cent method, phi, biserial r , point biserial r , and the discrimination index produce quite comparable reliability coefficients at all difficulty levels.

As the $U-L$ 27 per cent and phi have the advantage of computational ease and method reliability which is quite comparable to the more painstaking methods, it would appear that the $U-L$ 27 per cent method or phi should be acceptable for most situations. This latter conclusion is in essential agreement with the findings of Humphreys (10) and Flanagan (9) who have

investigated the use of these methods of item analysis from several other aspects.

From the intercorrelations between methods (Table 5), it can be observed that all of the methods of item analysis, with the exception of tetrachoric r and tetrachoric r corrected, are more closely related to point biserial r and the t -ratio than they are to biserial r . Further, point biserial r and the t -ratio come very close to being correlational identities for both semesters' examinations. Tetrachoric r yields the lowest intercorrelations of any method; however, the intercorrelations are substantially increased when tetrachoric r is corrected for non-median dichotomization. When the magnitudes of method intercorrelations are examined, it is evident that the point biserial coefficient and the t -ratio are the most representative methods of those computed with the U - L 27 per cent method following quite closely.

E. SUMMARY

To determine the effect of item difficulty on method reliability, the following methods of item analysis are examined: U - L 27 per cent, U - L 10 per cent, phi, tetrachoric r , tetrachoric r corrected for non-median dichotomization, biserial r , point biserial r , the t -ratio, and the discrimination index. The reliability of all methods is lowest for easiest items and highest for items of median difficulty. From the viewpoint of method reliability the U - L 10 per cent method, tetrachoric r , and tetrachoric r corrected for non-median dichotomization are sufficiently inferior to the other methods to make their use questionable. The other methods produce quite comparable reliabilities at all difficulty levels with ease of computation suggesting that the U - L 27 per cent method or phi should suffice for most situations. An overlooked but important factor in item analysis is the effect of non-normally distributed criterion scores on the magnitude of test item validity. All the methods examined are affected by this factor regardless of the specific assumptions concerning the variable.

REFERENCES

1. ADAMS, J. F. The effect of non-normally distributed criterion scores on item analysis techniques. *Educ. & Psychol. Meas.*, 1960, 20, (In Press).
2. ASCHENBRENNER, R. A table of values of the product moment coefficient of correlation as determined by the proportions of successes on the X -variable in each ten per cent tail of the Y -variable distribution. Iowa City: University Examinations Service of the State University of Iowa, 1949.
3. COOK, W. W. The measurement of general spelling ability involving controlled comparisons between techniques. *Univ. Iowa Stud. Educ.*, 1932, 16, No. 6.
4. DAVIDOFF, M. D., & GOHEEN, H. W. A table for the rapid determination of the tetrachoric correlation coefficient. *Psychometrika*, 1953, 18, 115-121.

5. DAVIS, F. B. Item analysis data: Their computation, interpretation, and use in test construction. *Harvard Educ. Papers*, 1946, No. 2.
6. EDWARDS, A. L. Statistical Methods for the Behavioral Sciences. New York: Rinehart, 1954.
7. FLANAGAN, J. C. General considerations in the selection of test items and a short method of estimating the product-moment coefficient from data at the tails of the distribution. *J. Educ. Psychol.*, 1939, 30, 674-680.
8. ———. A table of the values of the product-moment coefficient of correlation in a normal bivariate population corresponding to given proportions of successes. Pittsburg: American Institute for Research, 1950.
9. ———. The effectiveness of short methods for calculating correlation coefficients. *Psychol. Bull.*, 1952, 49, 342-348.
10. HUMPHREYS, L. G. Commonly used statistical procedures. In J. P. Guilford (Ed.), *Printed Classification Tests: Aviation Psychology Research Report No. 5*. Washington: Government Printing Office, 1947.
11. JURGENSEN, C. E. Table for determining phi coefficients. *Psychometrika*, 1947, 12, 17-29.
12. LONG, J. A., & SANDIFORD, P. The Validation of Test Items. Toronto: Department of Educational Research, University of Toronto, 1935.
13. NEFZGER, M. D., & DRASGOW, J. The needless assumption of normality in Pearson's *r*. *Amer. Psychol.*, 1957, 12, 623-625.
14. PERRY, N. C., KETTNER, N. W., HERTZKA, A. F., & BOUVIER, E. A. Estimating the tetrachoric correlation coefficient via I. a cosine-pi table and II. correction graphs for non-median dichotomization. (Technical Memorandum No. 2, Office of Naval Research Project NR 150-044, *Studies of Aptitudes of High Level Personnel*). Los Angeles: Department of Psychology, University of Southern California, 1953.
15. RICHARDSON, M. W. The relation between the difficulty and the differential validity of a test. *Psychometrika*, 1936, 1, 33-49.
16. RUCH, F. L. Psychology and Life. (4th ed.) New York: Scott, Foresman, 1953.
17. SMITH, M. The relationship between item validity and test validity. *Contributions to Education*, No. 621. New York: Bureau of Publications, Teachers College, Columbia University, 1934.
18. VERNON, P. E. Indices of item consistency and validity. *Brit. J. Psychol. (Stat. Sec.)*, 1948, 1, 152-156.
19. WALKER, H. M. Elementary Statistical Methods. New York: Holt, 1943.

Department of Psychology
 Temple University
 Philadelphia 22, Pennsylvania

CONSIDERATIONS INVOLVED IN OBTAINING AVERAGES OF CFF DATA*

Department of Psychology, Michigan State University

THOMAS M. NELSON, S. HOWARD BARTLEY, AND DORIS DEHARDT¹

A. INTRODUCTION

The decision to average data can be said to cut across divergencies in choice of problem material and analytic device. The average is a basic methodological topic and averaging a fundamental experimental operation.

Experimentalists average data to find out something about behavior. Decisions involving the average affect the design of an experiment and, thus, potentially the character of the collected data. Since the outcome of the experiment, the form of the data, will have not only to do with solving the problem posed, but will lead to attributing certain properties or functions to the organism, decisions about averaging are not necessarily trivial.

The problem of the average may be said to be that of providing a number that will adequately serve as a representative of a biological or personal process or serve to describe some fact presumed to bear upon the problem in question. One cannot, for this reason, assume that the average as sheer number and mathematical operation is the whole fact. A clear psychometric definition, though necessary, is insufficient by itself, except in special cases. The statistical operations describe only the tool and are not definitive of the chief concern.

Design of critical flicker experiments almost always involves some form of data averaging. Some decision ordinarily must be made regarding both "how many" and "which" of the responses are to constitute the averaging unit. The present study was undertaken to provide information bearing upon this problem. Our concern most specifically is with the effects various size averaging units have upon the value of the CFF.

B. OBSERVERS AND APPARATUS

A total of 30 observers, 18-57 years of age, were used. All met a visual acuity criterion of 20/30. Two were sophisticated and practiced observers.

* Received in the Editorial Office on February 2, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ The present study is a part of the work being done under Research Grant NSF-G-5821 from the National Science Foundation.

Eight were sophisticated and unpracticed. Twenty-one were unsophisticated and unpracticed. The definitions of these *O* classifications may be found in an earlier paper (25). Measurements indicated all had ample accommodation for the experimental conditions used. At the beginning of each experimental period, five minutes were spent in complete darkness. The *O*s served at the same time of the day when more than one experimental period was necessary.

The experiment was conducted in a darkened room. Stimulation was provided by means of an episcotister described at length elsewhere (4). Four target intensities were involved. Intensity was varied, using neutral density filters in approximately logarithmic steps from .13 (D4) to 1164.00 c/ft² (D0). Disks with one open sector and with pulse-to-cycle fractions (PCFs) of 29/30, 7/8, 3/4, 1/2, 1/4, 1/8, and 1/30 were alternately used. The target viewed subtended a visual angle of about 1.5 degrees and was entirely foveal. A chin rest fixed the head in position. The viewing distance was 30 inches.

C. PROCEDURE I

One *O* (who will be called *O*1) was sophisticated but unpracticed and served on three separate days. Each day she was instructed to make alternate ascending and descending observations. The ascending task was to report the disappearance of flicker in a target as the rate of stimulation was increased. The descending task was to report the emergence of target flicker when a high rate of stimulation was decreased. A PCF of 3/4 was used each day in conjunction with one of the lower three intensities. The first day used intensity D3, the second D2, and the last D4. The experimental periods were broken into "blocks" of 12 observations with 90 seconds allowed between each "block." The last 10 readings for each condition were used in the analysis of results. The first two responses were customarily discarded, since they are generally considered to have more to do with the observer's adjustment to the apparatus than with adjustments to the variable in question. In each experimental session, "blocks" of readings were taken until *O*1 reported that she was "too tired to go on." Introjection of this latter feature into the experimental procedure allows one to obtain the total number of non-fatigued responses *O*1 is capable of giving, and hence produces the *total* data an *E* would have available for averaging from this single occasion, whatever averaging procedure *E* might decide to adopt.

D. RESULTS I

This methodology produced 22 blocks of 12 readings each on Day 1, 24 on Day 2, and 30 on Day 3. The first two responses of each "block" were

discarded and arithmetic mean *CFFs* computed for each block of 10 trials on each day of observation. The most apparent feature of these data, is the general decreases in mean values as the number of observations is extended. The decrease is not, however, necessarily continuous over the whole series of observations, as the data for the second day increase initially very noticeably.

In order to present the data in a way that will more clearly show that progressive changes in the magnitude of the arithmetic mean occur as a function of the location of trials to be averaged in the sequence of trials,

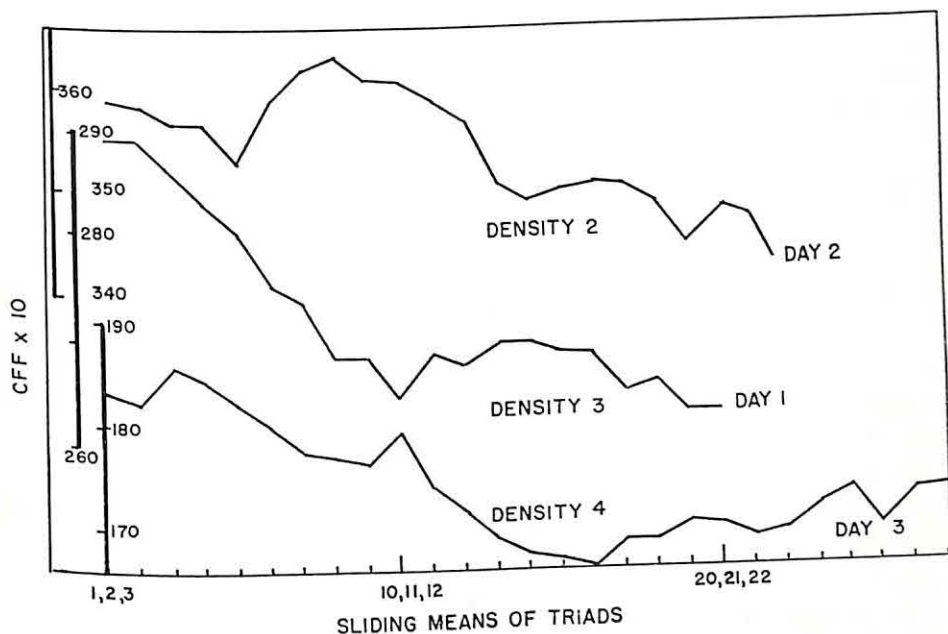


FIGURE 1

Curves showing *CFF* readings for a single *O*. Each curve represents a single sequence for a separate day. Each day the *O* was asked to continue readings till unable to continue. Stimulus intensity (labeled by filter density) is given in the text.

combined means are plotted in Figure 1. These averages represent averages of overlapping trial "blocks." That is, the mean of trial "blocks" 1, 2, 3, were found, then "blocks" 2, 3, 4, then 4, 5, 6, and so on. Each curve represents a single day. The general decreases are obvious graphically.

In order to statistically demonstrate that the location of averaged units affects the average, mean *CFFs* for Day 1 for Blocks 1, 11, and 22 were compared. It was hypothesized that the first, middle, and last blocks, when compared, would be of homogeneous variance and the mean differences statistically significant. Comparisons of Block 1 with 11 and Block 1 with 22

were statistically significant, with Block 1 significantly larger than Blocks 11 or 22. The mean comparison between Blocks 11 and 22 was not significant. All considered, the statistical analysis supports the impressions of the functions gained from visual inspection of curve form. The tabular and plotted means show a general decrease as the experiment progresses. The rates of intermittency of photic stimulation necessary to produce a fusion response change markedly as the number of observations increase from 1. This suggests that it is not the sheer number of trials averaged, but the place from within a series of responses that determines the stability of the mean value. Apparently a factor or factors are operating to change response as a given series of *CFF* observations is extended.

E. PROCEDURE II

We will now turn to an examination of samples from the first block of responses for a set of *O*s in order to ascertain whether responses in representative first "blocks" manifest analogous characteristics, even though on a smaller scale, to the entire repertory of responses when they are subjected to ordinary averaging procedures.

For the 29 *O*s, all 7 *PCF*s, but only the upper four target intensities, i.e., *D0*, *D1*, *D2*, *D3*, were used. For the practiced *O*s, 12 readings were taken over each of these 28 conditions and, in addition, *PCF*s of $1/4$ and $3/4$ were re-run over all four target intensities. Each practiced *O* served under 36 experimental conditions then, and there were a total of 72 units of data for the two practiced *O*s together. Four conditions of density were run each experimental period, i.e., *D3*, *D2*, *D1*, *D0*, respectively. The sophisticated *O*s were treated identically except that 12 additional readings were taken on the conditions of *PCF*s $1/4$, $3/4$, and $29/30$, so that a total of 240 units of data were taken for sophisticated *O*s. Six naïve *O*s received eight conditions at each of *PCF*s $1/4$ and $3/4$. Fifteen naïve *O*s received four conditions at two *PCF*s. The two *PCF*s for these 15 *O*s were combinations of *PCF*s $1/30$, $1/8$, $1/2$, $7/8$, $29/30$, and were chosen so that all five of these *PCF*s were presented six times. Each experimental period utilized the same four conditions of density as for the previous group of *O*s. Thus, a total of 216 conditions was presented to naïve *O*s. The number of units of data for these three groups of *O*s totaled 528. For all 29 *O*s, the last 10 readings for each experimental condition were used in the analysis of the results. The first two responses were always discarded for the reason previously stated.

F. RESULTS II

The data reported here are those used in a previous paper (25), but for another purpose. The 528 units of data from the 29 *O*s are treated in such a way that the results from all of these conditions can be given a common analysis, and so that the effects on the arithmetic mean of averaging various *numbers* of trials can be stated.

In order to arrive at a common means for expressing the results, the mean of the first 2, 4, 6, 8, and 10 observations was found for each of the 528 units of data. That is, five means were found, each computed from progressively more observations. This replicates situations that would occur in actual experimentation where a decision to make so many observations involves averaging all responses in the observational unit, be they 2, 4, 6, 8, etc., observations. Each of these five means was subtracted from the final mean (mean of all 10 observations), the last difference obviously being 0. The results of all 528 data units were, therefore, made comparable, since all pairs of scores were converted to relative differences in magnitude from the final mean of a condition. The absolute values of these differences from the final mean were then plotted for each of the 528 units, so that 528 graphs were obtained. From these plottings, eight general types of curves were observed. Prototypes of these eight graph types (called Types I-VIII) are reproduced in Figure 2.

The 8 curves were defined using the following observational criteria. All 528 graphs were examined briefly before the establishment of types began, so that a general impression of the various forms of curves could be obtained. Then the graphs were examined more carefully, one at a time, a type being specified every time a curve of a new form was encountered. A change in shape and/or slope in a graph from the previous graphs was tantamount to defining a new type. The height of the curve was not considered in the definition. The first graph examined defined a type, since it obviously was the first graph of its type to be observed. All prototypes were drawn as they were observed. The second graph defined a new type only if, upon examination, it was found to be of a different form from the first graph, and so on. Variation was permitted for no more than one point for each graph, and then only when the variation did not markedly alter the shape or slope of the general graph type. The variations which were permitted were defined after establishment of each prototype by the addition of dotted lines to the prototype, the result being that permissible variations for each type was specified and constant in use. This procedure was followed until all 528 graphs were dealt with. At the termination of classification,

the sample was discovered to contain the eight kinds of response patterns or "types" shown in Figure 2.

Once established, these "types" were investigated by relating them to known experimental variations. In order to do this, the frequencies of occurrence of each Type under the various conditions were ascertained. Since the variables of *PCF* and density were first considered, the number of graphs each *PCF*-intensity condition contributes to each category was deter-

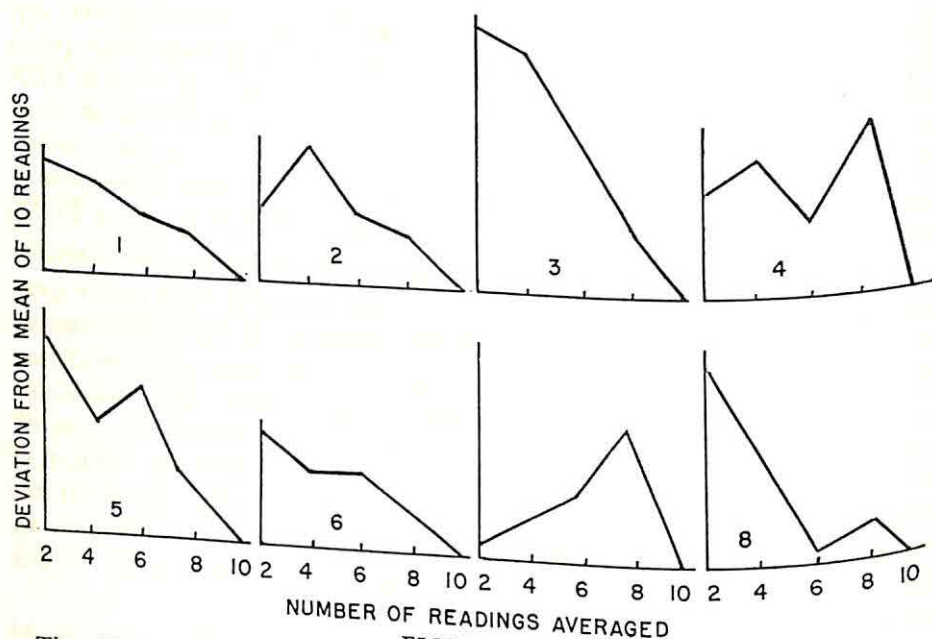


FIGURE 2

The eight response Types described in text. The various points on the curves represent successive (cumulative) means for 2, 4, 6, 8 readings in terms of their deviation from the mean of the total 10 readings. There were various numbers of cases found to be represented by these Types, as follows: Type 1, 29 cases; 2, 91; 3, 160; 4, 26; 5, 90; 6, 28; 7, 7; and 8, 97, making a total of 528 cases.

mined. Thus, a graph obtained for $PCF = 3/4$, $D = 2$ condition, which has the form of the Type III curve, results in a tally of $3/4 - 2$ for Type III, and so on. The results of this analysis are given in the legend of Figure 2.

It can be seen from Figure 2 that the bulk of the graphs were classified as Type II, III, V, or VIII. These types plus Type I all are decreasing functions and are typical of 88 per cent of the total units of data. A χ^2 was computed to test the hypothesis that number of graphs in each category was random, assuming equal theoretical frequencies (66 units per cell, i.e.)

for each of the eight categories. From this test, the hypothesis of randomness was rejected. Some categories apparently contain significantly greater numbers of the individual representative than others.

Two further χ^2 analyses were done on the categories taken separately. The purpose of the first analysis was to determine whether any graph Type (Type I-VIII) includes greater than a random number of any one or more of the seven *PCFs*. If such were the case, it would suggest that certain of the sequential features in response are tied to these changes in nature of the photic impingement. To test this, eight χ^2 's were computed, one for each Type, the hypothesis being that the number of graphs occurring within any of the eight categories will correspond to the proportion of each *PCF* in the total population of 528. These χ^2 's were not significant at even the 20 per cent level of confidence. Thus, the hypothesis of random occurrence of the *PCF* variable within Types cannot be rejected. These results indicate that no classification is unique to a given *PCF* or group of *PCFs*.

The second analysis was done to determine whether one or more densities occur disproportionately often within any of the Types. If so, it would mean that intensity is closely associated with the unique or characteristic response form of a particular Type. This is an important test since some might assume the response forms to reflect physiological impairment. If so, one would expect high intensities to behave differently than low intensities in this regard. Again, eight χ^2 's were computed, the hypothesis being that the number of graphs for each density within a given category is random. None of the eight χ^2 's was significant. The hypothesis is not rejected. There is no evidence that differences in intensity are peculiar or characteristic of any of the response Types.

Since 14 *O*s in this study were presented several conditions twice (eight were presented *PCFs* 1/4 and 3/4, and six had *PCFs* 1/4, 3/4, and 29/30 on two separate sessions), these data were analyzed to determine whether the graph Types observed for a single *O* would vary as a function of amount of practice. It might be argued that the eight Types of curves are learning curves, particularly since many of them were constantly decreasing functions. If this is so, then one would expect that curves obtained during the second session (for the same conditions) would differ from those for the first session. A total of 136 conditions were repeated or re-run. A χ^2 analysis was made on the distribution of graph Types of these 136 conditions for the two sessions. Theoretical frequencies were assumed to be the mean of the observed frequencies at each graph Type. The obtained $\chi^2 = 2.74$, $df = 7$, was highly insignificant, indicating that the distribution of graphs for Session

2 did not differ from that of Session 1, or that the graph Types obtained with the conditions used in this study are not dependent on amount of practice with these conditions.

A χ^2 was also computed to test the hypothesis that distributions of *O*s for the three classifications ("naïve," "sophisticated," and "practiced") were random across the eight graph Types. The theoretical frequencies were, therefore, based on the assumption of equal probability of occurrence of an *O*'s response falling in one of the eight graph Types (with corrections for the differences in *N* for the three *O* groups and differences in *N* for the eight Types). A $\chi^2 = 16.13$, $df = 14$, was obtained which was highly insignificant, indicating that distributions of graphs for *O*s from the three classifications across Types was random. The class of *O* is apparently not an important variable.

G. DISCUSSION

The data from these 29 *O*s, as well as for *O*1, demonstrates that an average of *CFF* responses is affected by serial operations used in experimentation. The data of the 29 *O*s show these operations in *CFF* to affect averages taken from even relatively small units of data. Using 528 blocks of 10 responses, serial effects are almost invariably present. These effects are initiated at the very outset of observation so that a block of 10 observations is no more free of these than is a block of 200. The magnitude of a *CFF* mean must be considered to be dependent upon both the *number* of responses and upon the *position* of the average unit within the range of possible serial positions. The data for *O*1 showed decreasing values throughout the whole range of possible observations under non-fatigued conditions.

The changes within the data for the 29 *O*s are characteristic enough to be classified into Types. Eight Types of serial responses were discoverable when shape was considered. This suggests that the serial factor or factors in operation do not lead to rigid stereotyping of response. In addition, although ordinarily both the curves representing the individuals' responses and those representing response Types have continuously changing slopes, the slope fluctuates in some of the cases. The presence of multiple Types, the differences between Types and the fluctuations evident in the data of *O*1, especially the reversal of Day 2, suggest that these general changes are not dependent upon the *continuous* propagation of some spurious influence over an entire series of judgments, as one writer (23) has suggested might be the case.

Attempts to ascertain the nature of response variability by relating Types to known experimental variables resulted in statistically non-significant levels

of confidence. This was the case even though 28 different instrumental conditions of observation were employed in Procedure B. Apparently the functions are not secondary effects of the levels of photic intensity or *PCFs* used. The fact that the observers do not show different serial effects at the extreme levels of photic intensity also seems to rule out an explanation of the serial phenomena on the basis of progressive physiological impairment (2), or so-called "physiological fatigue."

Re-run trials on the same individuals under the same conditions did not alter the character of the response patterns. Thus, "learning" seems definitely precluded as a principal contributor. This was not unexpected (34).

Explanation of the results in terms of peculiarities in observer population seems ruled out. "Naïve," "sophisticated," and "practiced" *O*s were all used. The functions of the eight Types could not be attributed to response peculiarities between these populations of observers providing the responses. While the fundamental difference between the three classes of *O*s is known to be one of variability (25), and the Types are representative, in a sense, of forms fluctuation in response may take, the differences produced by use of three observer populations was shown here not to be manifested as some particular kind of serial response Type.

The discovery that procedures requiring serial response produce an effect is certainly nothing new. Since Urbantschitsch first reported changes in auditory thresholds upon serial stimulation 85 years ago, "serial effects," "non-independence of successive responses," "ordinal effects," etc., have frequently been made the topic of inquiry (7, 10, 11, 13, 22, 23, 26, 33, 34, 38). Unfortunately, the numerous previous investigations into serial response data cannot be said to provide much information bearing on the present case. First of all, none of them is concerned with *CFF* situations. Second, most of the research reported is tied to operations affecting fluctuations at absolute threshold. Third, the case in which there is a continuous variable under the control of the observer is not ordinarily considered. Lastly, there has been a tendency, most marked in recent research reporting, to select out variables such as "prior response" and "prior stimulus conditions" for study, to provide experimental conditions to maximize their effect, and then to conceptualize these within theoretical frameworks, often "mathematical models," quite far removed from the concrete task performance and biology of the observer. Such results, conclusions, and theories could not be made to bear upon the methodological problem outlined to any advantage. Hence, one is led to seek an understanding of the natural basis of the progressive changes in response elsewhere.

It is possible that the changes in *CFF* values are associated with disorganizations eventuating in fatigue. The fact that we were able to produce the progressive decrease in three cases (*O1* data) when the performance was known to eventuate in fatigue (by fatigue we mean those processes associated with tiredness) (2), suggests this to be the case. Further, most of the Types discovered in the sample of data from the 29 *O*s show continuous functions similar to these known fatigue curves. One form of disorganization leading to these effects might even be conjectured, since certain other independent experimental findings seem to fit the present research picture.

Various studies concerning the functioning of the accommodative mechanism in serial visual performances have shown changes in visual resolution to occur in many situations demanding prolonged focusing (5, 16, 18, 19, 20, 21). These fluctuations in the accommodative mechanism fit the definition of disorganization (2) since they are more in the nature of a temporary lack of coordination than they are of breakdown in local tissue (16). In the cases reported, sustained observation produced increases and decreases of accommodative amplitude, and, most significantly, sudden accommodative shifts. Shifting accommodation could produce the serial response types noted since it would shift the resolution of residual target flicker in a *CFF* experiment. Also, an interpretation based upon disorganization of the accommodative mechanism is bolstered by both the nature of the response situation and the verbal reports given by our *O*s. Our experimental situation was highly restrictive. Restrictive situations are those thought most likely to produce the conflict resulting in disorganization of peripheral mechanisms (2). The head of *O* was fixed in position, and for a rather prolonged period he intently viewed a constantly but subtly changing target falling on less than two degrees of the fovea. The introspective reports from the *O*s led one to infer that the serial procedure was producing oculomotor difficulties. Introspective data given by the 29 *O*s, as well as the data of *O1*, are replete with reports of discomfort and inability, strongly suggesting that accommodation was difficult to maintain. Our discovery of response types is reminiscent of Howe's (18, 19, 20) findings in another accommodative task.

While it can be conjectured that the features described reflect the operation of a primary variable, it cannot be overemphasized that the serial effects themselves are matters of fact differing in no way from other variables in response. As such, they must be controlled for by experimenters just as are *PCF*, photic intensity, spectral composition, area, etc. Serial response is not a factor to be treated merely as a formal item within experimental

design, although presumably it might be so described. In particular, it will clearly not do to base methodological decisions upon the ordinary premise to the effect that there is some best number or range of trials (that a "true" mean *CFF* value exists), or/and that, within very broad limits, the greater the number of readings taken per session the better. Like it or not, averaging operations are beholden to considerations apart from the mechanism of mathematical analysis. The problem is to devise a methodology allowing for the production of *CFF* data that will control for serial effects and thus provide more reliable comparisons to be made between data from various experimental conditions than is presently possible.

Initially, it was anticipated that the effects from serial operations would be manifested by general disjunctions in performance. We hoped to find abruptions in performance as a function of number of trials that would allow the making of some set of absolute recommendations, such as to the number of trials to use. Since disjunctions did not occur, other alternatives for experimental economy must be considered.

One alternative of control would be derivation of a factor designed to nullify the serial effects for individual *O*s. Such a factor would be best if designed to be useful with various sized units of data. If individual *O*s are stable enough in their serial tendencies to establish a corrective equation under one set of instrumental conditions, it may be possible that the equation can serve as a general corrective device for any one *O* under any condition, since shifts in instrumental conditions seem not to affect the relative frequency of any Type occurring. An idea of a single general corrective factor for all populations might also be entertained except that populations of "naïve" *O*s are not particularly desirable (25) and large populations of the most desirable *O*s, the practiced *O*s, are not ordinarily available.

Another alternative is that of taking a very few *CFF* readings each day and basing the average upon a great many occasions of observation. Let us say that, instead of 12 observations at one sitting, four were taken on each of five days to arrive at a block of 10 responses for averaging (the first two in each case always being discards). If, as in this experiment, 28 conditions are used, 140 experimental sessions per *O* would have to be scheduled instead of the seven of the current methodology. Practical considerations alone will probably rule this out for most experimenters. Moreover, other factors might be working against experimental control at the same time. Hecht, Hendley, and Ross (14) established that even rather brief increases in natural out-of-doors illumination alter the quality of visual performance in a way that is measurable for days afterwards. Considerable numbers of

biological periodicities and long term changes are also known (2), and even non-photoc shifts in weather may have to be reckoned with as variables affecting human visual performance (27).

One possibility for control might seem to lie in randomization of the serial position of various instrumental conditions. It can be argued that the most simple control is to assure that all instrumental conditions have all serial positions across an experiment. However, because of retinal adaptation, it will not do to use considerable instrumental variations in *intensity* in indiscriminate order, as one may in the case of certain other variables. Twenty minutes for dark adaptation between responses in which weak photic pulses follow those of high intensity cannot ordinarily be used. Even if such amounts of time were taken, the rather high levels of illuminations occurring prior to the weak in some cases can be expected to affect subsequent performance after longer intervals (14) and hence introduce complications into the experiment.² While a statistical analysis in conjunction with randomization might conceivably show the presence of a significant difference between instrumental conditions, it is cumbersome, if not impossible, to satisfy oneself concerning the magnitude of the effect that changes in instrumental conditions provide. Randomizations act to mask the effects of the experimental variables as well as those *E* is attempting to control.

It is doubtful whether any fixed set of common operations can be stipulated that will cover any and all *CFF* situations. The aim of securing stable phenomena seems best served by a general policy whose particulars can be tailored to fit the specific occasion. With respect to policy, three things can perhaps be said.

First, the same number of observations should be made under each experimental condition used. It will be recalled that Types were not differently distributed when instrumental conditions were changed, and that these conditions occurred in four serial positions at a time (see methodology, Procedure B). Thus, when the response units are kept of equal length, at least within the range of serial operations we investigated, one can be fairly well assured that the serial effects are equivalent from response block to block.

Second, it should be recognized that, although many lengths of response series are possible, it is best that all experimenters use the same sized blocks if experimental findings are to be compared. A preference for one length of response series over another is probably best made by considering variability. Variability can probably be kept down by following a rule that the fewer

² Control by means of counterbalanced conditions is hence not possible either.

responses taken at one time the better. The rule might be amended, where multiple experimental conditions are the case, to say "the greater the variety of experimental conditions, the fewer the observations per condition per session." Our experience and data from re-trial series suggests that, when reasonable precautions are taken with regard to general health, scheduling time of day for observations, etc., less variability can be expected from an *O* from one day to another than occurs during a single period of prolonged observation.

Third, when unusual series of observations are necessary, the operational characteristics of the series should be clearly described, the reasons for using it, and its possible effects upon the average stated. This will allow experimenters to control for the serial effects introduced if related experimentation is subsequently undertaken.

When evaluating these suggestions, however, one should keep to the foreground the fact that suggestions concerning methodological practices are only so sound as the best knowledge concerning a phenomenon allow. Facts discovered about a variable in one group of experiments serves to better define later experiments involving this variable in an indirect way, for control of a variable demands that one know its mode of operation over a considerable range of conditions. The effects of serial operations is an unknown in human performance to which experimental procedures must be applied, and concerning which new information is to be gained. This reason alone should make serial response a variable warranting more general experimental attention and research.

H. SUMMARY

CFF data from 29 *O*s in one type of experiment and 1 *O* in another were examined to determine whether they displayed characteristics which would provide for generalizations in regard to the procedure of averaging. While principles in regard to averaging of data were the primary consideration, the determination of the fewest readings to be obtained for any single condition was attempted. No blanket rule in regard to numbers of readings to be averaged was formulated, but the data provided the basis for what was considered to be a helpful discussion of the subject.

REFERENCES

1. ABELSON, R. P. Spectral analysis and the study of individual differences in the performance of routine repetitive tasks. Princeton, N. J., *Educ. Test. Serv.*, 1953.
2. BARTLEY, S. H., & CHUTE, E. Fatigue and Impairment in Man. New York: McGraw-Hill, 1947.

3. BARTLEY, S. H. Central mechanisms in vision. In Field, J., Magoun, H. W., & Hall, V. G., (Eds.), *Handbook of Physiology, Sec. 1, Neurophysiology, Vol. I.* Washington, D. C.: Amer. Physiol. Soc., 1959. (Chap. 30, Pp. 713-741.)
4. BARTLEY, S. H., & NELSON, T. M. Equivalence of various pulse-to-cycle fractions in producing critical flicker frequency. *J. Opt. Soc. Amer.*, 1960, in press.
5. BERENS, C., & STARK, E. K. Fatigue of accommodation. *Am. J. Ophth.*, 1932, 15, 527-542.
6. COLLIER, G. Probability of response and inter-trial association as functions of monocular and binocular stimulation. *J. Exp. Psychol.*, 1954, 47, 75-83.
7. COLLIER, G., & VERPLANCK, W. S. Nonindependence of successive responses at the visual threshold as a function of interpolated stimuli. *J. Exp. Psychol.*, 1958, 55, 429-437.
8. COTTON, J. W., & VERPLANCK, W. S. The dependence of frequencies of seeing on procedural variables: II. Procedure of termination series of intensity-ordered stimuli. *J. Gen. Psychol.*, 1955, 53, 49-57.
9. COTTON, J. W., & VERPLANCK, W. S. The dependence of frequencies of seeing on procedural variables: III. The time-interval between successive stimuli. *J. Gen. Psychol.*, 1955, 53, 59-66.
10. DAY, W. Serial non-randomness in auditory thresholds as a function of inter-stimulus interval. *Amer. J. Psychol.*, 1956, 69, 387-394.
11. FERNBERGER, S. W. Independence of judgments within the series for the method of constant stimuli. *J. Exp. Psychol.*, 1920, 3, 126-150.
12. FLYNN, J. P., JEROME, E. A., & MOODY, J. A. The interserial dependence of psychophysical responses. *Amer. Psychol.*, 1952, 7, 281-282. (Abstract)
13. GUILFORD, J. P. "Fluctuations of attention" with weak visual stimuli. *Amer. J. Psychol.*, 1927, 38, 534-583.
14. HECHT, S., HENDLEY, C. D., & ROSS, S. Influence of exposure to intense sunlight on subsequent night vision. M and S Res. Proj. No. X-422 (AV-233-W), 26 April, 1945.
15. HELSON, H. Adaptation-level as frame of reference for prediction of psychological data. *Amer. J. Psychol.*, 1947, 60, 1-29.
16. HOFSTETTER, H. W. An ergographic analysis of fatigue of accommodation. *Am. J. Optom. Arch. Am. Acad. Optom.*, 1943, 20, 115-135.
17. HOLLINGWORTH, H. L. The central tendency of judgment. *J. Phil. Psychol. Sci. Meth.*, 1910, 7, 461-468.
18. HOWE, L. The fatigue of accommodation. *J.A.M.A.*, 1916, 67, 100-104.
19. ———. The fatigue of accommodation as registered by the ergograph. *Tr. Sect. Ophth., A.M.A.*, 1916, 130-142.
20. ———. Registration of fatigue of accommodation. *Tr. Am. Ophth. Soc.*, 1917, 15, 145-153.
21. LANCASTER, W. B., & WILLIAMS, E. R. New light on the theory of accommodation, with practical applications. *Tr. Am. Acad. Ophth.*, 1914, 170-195.
22. LONG, L. A study of the effect of preceding stimuli upon the judgment of auditory intensities. *Arch. Psychol.*, No. 209.
23. MCGILL, W. J. Serial effects in auditory threshold judgment. *J. Exp. Psychol.*, 1957, 53, 297-303.
24. MICHELS, W. C., & HELSON, H. A quantitative theory of time-order effects. *Amer. J. Psychol.*, 1954, 67, 327-334.
25. NELSON, T. M., BARTLEY, S. H., & DEHARDT, D. A comparison of variability of three sorts of observers in a sensory experiment. *J. of Psychol.*, 1960, 49, 3-11.
26. PARDUCCI, A. An adaptation-level analysis of ordinal effects in judgment. *J. Exp. Psychol.*, 1959, 58, 239-247.

27. PETERSON, W. F., *et al.* The patient and the weather. Ann Arbor, Mich.: Edwards, 1934-1936.
28. SCHAFER, J. H. Influence of preceding item in measurements of the noise-masked threshold by a method of constant stimuli. *J. Exp. Psychol.*, 1950, **40**, 365-371.
29. SENDERS, V. L. Are successive psychophysical judgments in a series independent? *Amer. Psychol.*, 1952, **7**, 282. (Abstract)
30. SENDERS, V. L., & SOWARDS, A. Analysis of response sequences in the setting of a psychological experiment. *Amer. J. Psychol.*, 1952, **65**, 358-374.
31. SENDERS, V. L. Further analysis of response sequences in the setting of a psychophysical experiment. *Amer. J. Psychol.*, 1953, **66**, 215-228.
32. STEVENS, S. S. On the psychophysical law. *Psychol. Rev.*, 1957, **64**, 153-181.
33. VERPLANCK, W. S., COLLIER, G. H., & COTTON, J. W. Nonindependence of successive responses in measurements of the visual threshold. *J. Exp. Psychol.*, 1952, **42**, 273-282.
34. VERPLANCK, W. S., COTTON, J. W., & COLLIER, G. H. Previous training as a determinant of response dependency at the threshold. *J. Exp. Psychol.*, 1953, **46**, 10-14.
35. VERPLANCK, W. S., & COTTON, J. W. The dependence of seeing on procedural variables: I. Direction and length of series of intensity-ordered stimuli. *J. Gen. Psychol.*, 1955, **53**, 37-47.
36. VERPLANCK, W. S., & BLOUGH, D. S. Randomized stimuli and the nonindependence of successive responses at the visual threshold. *J. Gen. Psychol.*, 1958, **59**, 263-272.
37. WEISS, B., COLEMAN, P. D., & GREEN, R. F. Model for time-ordered dependencies. *J. Exp. Psychol.*, 1955, **50**, 237-244.
38. WERTHEIMER, M. An investigation of the "randomness" of threshold measurements. *J. Exp. Psychol.*, 1953, **45**, 294-303.
39. WILLINGHAM, W. W. Interdependence of successive affective judgments. *Amer. J. Psychol.*, 1959, **72**, 103-106.

Department of Psychology
Michigan State University
East Lansing, Michigan

MAXIMUM I_t FOR ABSOLUTE IDENTIFICATION OF CUTANEOUS ELECTRICAL INTENSITY LEVEL*

Psychology Division, U.S. Army Medical Research Laboratory

GLENN R. HAWKES AND JOEL S. WARM

A. INTRODUCTION

Information theory permits a measure of the accuracy of performance when making choices among equally likely alternatives of unidimensional stimuli. From a limited set of stimulus and response alternatives, the subject identifies as many stimuli as can be encompassed by a subjective scale. Miller (10) states that only seven alternatives (2.6 bits) may be the limit of this scale for most stimuli, this limit being the Channel Capacity.

Alluisi (1) distinguishes between Miller's concept of man's general Channel Capacity, and the channel capacity for a specific task. The former is described as the maximum amount of information that a subject can give about a set of stimuli regardless of task, the latter as the highest value possible for a specific task. The channel capacity determined for a specific task has been found to be dependent upon a number of experimental conditions. Frick (5), e.g., points out the necessity of having equal discriminability between stimuli in the series presented, and an equal number of *S* and *R* categories. Further, knowledge of results (feedback) has been found to increase performance in a wide variety of tasks (3, 4).

Alluisi lists the above variables, and also the extent of the physical range of stimulation used, as being primary experimental variables in the measurement of absolute identifications. According to his review, the physical range of stimulation appears to be the largest single determinant of information transmitted (within a given stimulus dimension). In general, the greater the stimulus range the greater the amount of information transmitted (I_t). He goes on to state that unless the variables of equal stimulus discriminability, equal number of *S* and *R* categories, knowledge of results, and physical range of stimulation used are optimal when determining the channel capacity for a particular function, the resulting value will represent only the maximum *S-R* correlation that the subject can obtain under the particular conditions employed, and will usually be less than the channel capacity for that task.

* Received in the Editorial Office on February 2, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

The Channel Capacity, or subjective scale, is not a measure of differential discrimination limits, and is related to discriminability only by the fact that the stimuli presented for identification must be at least one *j n d* apart along a particular dimension. Extension of the physical range of stimulation employed may indeed affect I_t , probably by permitting a wider spacing of the stimuli, thereby making the stimuli more closely approximate 100 per cent *j n d* spacing on the dimension studied. It is possible, of course, that additional stimulus relationships are produced by such a procedure, as in the "octave relationship" in judgments of auditory pitch.

Knowledge of results may also increase I_t for a particular experiment. Aside from any additional cues furnished by feedback (transitory learning effects), however, any increase in I_t would probably be due to "anchoring" effects. Immediate feedback to the subject would tend to prevent shifts in adaptation-level and provide repeated anchoring of the subjective scale.

Ability to make absolute identifications of current intensity levels was investigated by Hawkes (8) in an attempt to determine the usefulness of intensity as a primary cue in an electrical cutaneous communication system (cf. 7). Hawkes presented stimuli within an intensity range selected to avoid pain for most subjects, at frequencies of 100 and 1500 cps. Results indicated that the "amount of I_t by the average *O*" (cf. 1, p. 98) reached a maximum of 1.18 bits for those receiving stimulation at 100 cps, and 1.35 bits for those stimulated at 1500 cps. A significantly greater amount of information was found to be transmitted by the "average *O*" of the latter group.

Under the experimental conditions employed, however, the span of absolute judgment for current intensity appeared to be attained by the use of only three or four levels, irrespective of the stimulus frequency. From these data, it was concluded that the use of level of intensity as a primary cue in an electrical cutaneous communication system would be restricted to only two levels in a system requiring perfect accuracy, and to three levels in a system that could tolerate some error in order to increase information transmitted.

In the above study, the factors of stimulus spacing, and the number of response categories in relation to the number of stimulus categories, were optimally controlled by the provision of equal spacing by apparent subjective magnitude of sensation, and the use of response categories equal to the number of stimulus categories. However, the physical range of intensities studied was restricted to that between the vibration absolute threshold and twice the amount of current necessary to reach the absolute threshold for the individual observer in a given session, this range tending to avoid pain. Further, because it typically is not possible in communication systems, no

information as to the accuracy of his performance was afforded the subject during the testing session.

Hawkes and Warm (9) determined the tolerance limit (*TL*) for alternating current stimulation of the skin, and the reliability of the reports of these limits. Subjects were able to report the *TL* as reliably as the vibration or pain *RL* and evidenced no undue emotional reaction incident to the use of stimulus intensities which evoked pain. These data indicate that alternating current applied to the skin for communication purposes may feasibly include current values approaching the *TL*.

The present study was designed to investigate the effects of extension of the physical range of intensities close to the *TL*, and the provision of feedback information, on the subject's ability to absolutely identify current intensity levels. This procedure should provide a closer approximation of the channel capacity for the specific task involved, as defined by Alluisi, and of the general Channel Capacity for such identifications, as defined by Miller.

B. EXPERIMENTAL

1. *Subjects*

Ten subjects were used in this experiment, all of whom had participated in previous research involving the absolute identification of levels of current intensity.

2. *Apparatus*

The apparatus was the same as that of a previous report (8).

3. *Procedure*

The method of "magnitude estimation" described by Stevens (14) was employed to obtain data necessary for subjectively equal stimulus spacing (cf. 8). After preliminary investigation of the tolerance limits of the subjects, 20 stimulus intensities, progressing in intervals of 10 from 110 to 300 per cent of the current value necessary to reach the subject's *RL* were selected for judgment. For this series of judgments, 160 per cent of the *RL* current value was designated as the standard and assigned a numerical value of "10." The 20 stimulus values were presented in a different random order for each subject at a frequency of 1500 cps.

On the basis of the above results, four sets of stimuli consisting of three, four, five, and six current intensity levels were selected for absolute identification. Stimuli of each set were equally spaced in terms of apparent subjective magnitude of sensation.

In order to facilitate comparison with the previous study, after participation in the sessions described above, subjects also made absolute identifications of six intensity levels with a physical intensity range identical to the previous experiment. Current values for this six-category set were determined from Hawkes' magnitude estimation results (8). All stimuli of this experiment were presented at a frequency of 1500 cps.

The subjects were divided into two matched groups on the basis of previous performance. Subjects in Group 1 made absolute identifications of current intensity levels as in the prior experiment, with the stimulus values listed above, and were informed of the correct value of the stimulus after making each judgment. Subjects of Group 2 received identical current intensity levels for identification, but were not afforded immediate knowledge of results.

Subjects made 100 identifications per session, with a different number of current intensity levels being used in each session. All subjects received three levels in the first session, four levels in the second session, etc. Duration of a stimulus presented for absolute identification was 500 msec.

C. RESULTS

The stimulus intensity range was chosen so as to be always less than the *TL* value of the observers. The mean *TL* to *RL* ratio for the subjects was 4.38, $\sigma = 1.18$. A physical intensity range of 1.00 to 3.00 of the *RL* current value was selected.

Median estimates of apparent magnitude of sensation, plotted in log-log coördinates, are shown in Figure 1. The slope of the straight line which best fitted these data was 4.5.

The percentage of correct absolute identifications indicated that subjects in the present study were able to absolutely identify three current intensity levels with better than 95 per cent accuracy, and that subjects afforded knowledge of results made fewer errors than those without such knowledge. Both groups of this experiment tended to make more correct identifications than in the previous study.

I_t as a function of the number of current intensity levels for both the previous and the present experiment are plotted in Figure 2. Analysis of I_t was made according to a model detailed by Garner and Hake (6).

Although stimulus information increases with the number of intensity levels used, I_t in the previous study was little greater with the use of 4, 5, or 6 levels as opposed to the use of three such levels.

Accuracy of group matching on the basis of previous performance is

evidenced by the close relationship between the restricted range performance curves plotted in Figure 2. A t -test (for correlated samples) of the significance of the difference between the "amount of I_t by the average O " of each group under the restricted range condition indicates that the groups do not differ significantly ($df = 4$, $t = .10$, $p > .90$).

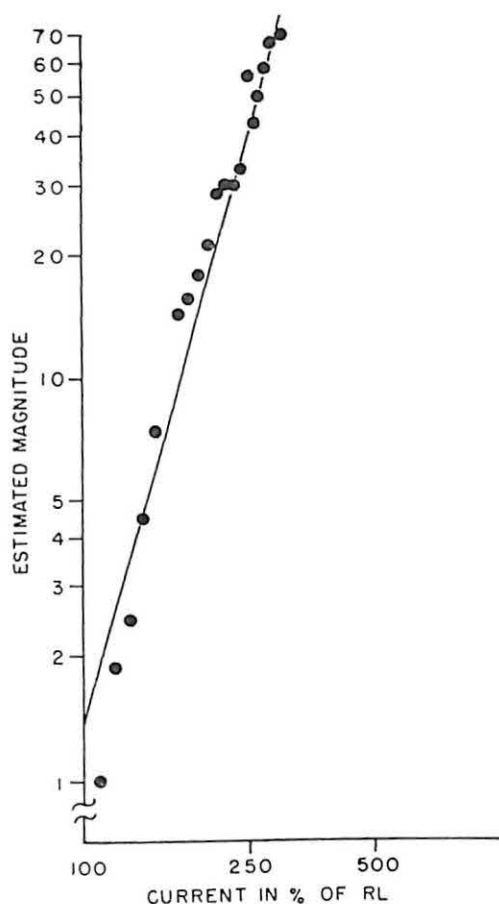


FIGURE 1
EFFECT OF CURRENT INTENSITY ON APPARENT SUBJECTIVE MAGNITUDE OF SENSATION

The effects of extension of the physical range of intensity of stimulation, and the provision of feedback information, on performance were evaluated by the use of Student's t for correlated samples. The "amount of I_t by the average O " of both groups of the present study was significantly greater with the use of a physical range of intensities close to the TL of the subject than with a physical intensity range restricted to a level below the pain thresh-

hold ($t = 6.39$, $p < .01$ for Group 1; $t = 3.27$, $p < .05$ for Group 2). Although Group 1 (extended range with feedback) appears consistently to transmit more information than Group 2 (without feedback), the mean difference in I_t by the average O of the two groups is not sufficiently large to permit rejection of the null hypothesis ($t = 1.94$, $p > .05$).

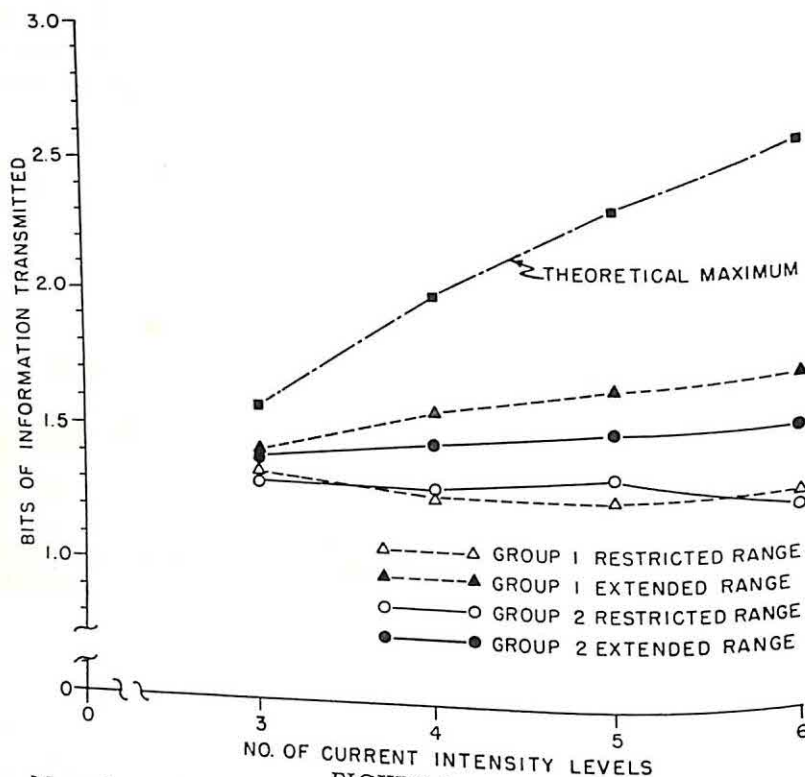


FIGURE 2
MEAN I_t VIA ABSOLUTE IDENTIFICATION OF CURRENT INTENSITY LEVEL

The relationship of response accuracy to the number of alternatives available to the subject (degree of uncertainty of the subject) is plotted in Figure 3. In each case, the function appears to be linear. The slope of the best fitting straight line for the data of the previous experiment is 1.03, for the condition of extended intensity range without feedback is .94, and for that of extended range with feedback is .72.

D. DISCUSSION AND CONCLUSIONS

Based on the literature reviewed in the Introduction, increased I_t by absolute identifications of current intensity levels might be expected to result

from: (a) extension of the physical intensity range used for stimulation, or (b) the provision of immediate knowledge of results. The data of this study indicate that a significant increase in the amount of I_t resulted from extension of the physical intensity range. The provision of immediate knowledge of results failed to produce a statistically significant increase in I_t ; however, a trend in the expected direction is indicated. Subjects afforded knowledge of results made fewer errors and transmitted more information than those without such knowledge. Further, in terms of the magnitude

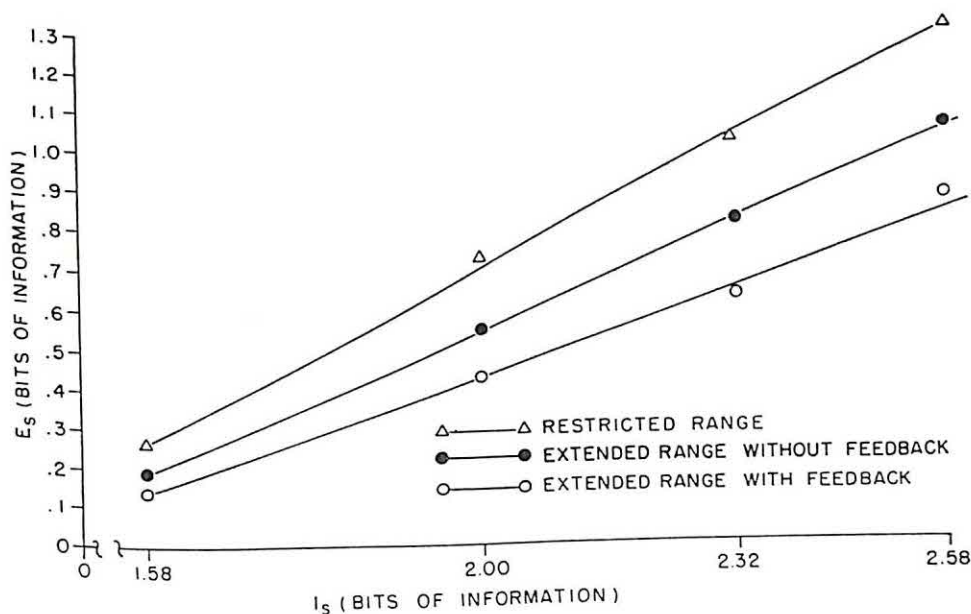


FIGURE 3
THE RELATIONSHIP OF SUBJECT UNCERTAINTY TO THE AMOUNT OF
STIMULUS INFORMATION

of perceptual error, subjects receiving knowledge of results made fewer "two-place errors" (identifying "2" as "4," "5" as a "3," etc.) than those receiving no feedback information (the with-knowledge group made 3 per cent errors of this type, and the without-knowledge group made 8 per cent such errors). These results are not inconsistent with the conclusion reached by Alluisi (1), viz. that the stimulus range appears to be the largest single determinant of I_t within a given dimension.

The maximum amount of I_t by the "average O" under the "optimal conditions of experimentation" employed in this study, was 1.72 bits, a value somewhat greater than the maximum of 1.31 bits reported previously

(8). The present data (see Figure 2) indicate that the span of absolute judgment of current intensity is attained by the use of four or five levels. Although the value is below the span of seven (2.6 bits) proposed by Miller (10) as the usual limit for absolute judgments of unidimensional stimuli, it is in accord with what might have been expected from results of studies of mechanical cutaneous vibration (7).

The data of the present study indicate that the apparent magnitude of sensation increased as a function of stimulus intensity at a rate somewhat greater than that reported in a previous experiment (8). The slope of the line which fitted the present data was 4.5, a value agreeing, however, with the prediction of Stevens (15, p. 204). The mean *TL:RL* ratio obtained for the present subjects also is in accord with the ratio reported by Hawkes and Warm (9).

Inasmuch as the present subjects also participated in a previous study of a similar nature, it is possible that some of the present improvement in performance (cf. 8) may be due to additional experience. It is doubtful, however, that much of the performance improvement in the present instance may be so explained. Examination of Hawkes' data (results of subjects receiving stimulation at 1500 cps) with respect to performance at three or four intensity levels of subjects identifying two intensity levels in the first session vs. those identifying five intensity levels in the first session, reveals no significant differences in I_t (t for three levels = .13, $p > .90$; for four levels $t = .17$, $p > .80$). Further, subjects made identifications of six intensity levels under the conditions of the previous report (restricted stimulus intensity range) after completion of the sessions of the present study. Mean I_t by the "average O " with this additional experience was no greater than their previous performance.

The relationship of the degree of uncertainty of the subject and the amount of information (number of alternatives) in the stimulus situation has been described by Weaver (13, pp. 108-109). A linear relationship between the degree of uncertainty of the subject (I_t to the subject) and response accuracy (I_t by the subject) has been demonstrated in the present investigation, consistent with studies in other modalities (2, 5, 11, 12). Although it appears that extension of the physical range of stimulation and the provision of immediate knowledge of results reduce the rate of increase in error with respect to an increase in the number of response alternatives, the amount of error related to the amount of stimulus information remains quite large when using more than three intensity levels.

The relatively small span of absolute judgment for current intensity

levels, and the relatively rapid rate of increase in perceptual error with additional uncertainty of the subject, indicate that the use of intensity as a cue for signaling purposes would be limited to three levels in a communication system requiring a high level of accuracy, and to four levels in a system capable of tolerating some error in order to maximize information transmitted.

E. SUMMARY

The effects on I_t of extension of the physical range of stimulus intensities, and the provision to the observers of immediate knowledge of results were investigated. The former resulted in a significant increase in I_t , the latter tended to increase I_t . Under optimum conditions, the use of intensity level as a cue for signaling purposes would be restricted to three levels in a communication system requiring a high level of accuracy, with four levels useful in a system capable of tolerating some error in order to maximize information transmitted.

REFERENCES

1. ALLUISE, E. A. Conditions affecting the amount of information in absolute judgments. *Psychol. Rev.*, 1957, **64**, 97-103.
2. ARCHER, E. J. Identification of visual patterns as a function of information load. *J. Exp. Psychol.*, 1954, **48**, 313-317.
3. ARPS, G. F. Work without knowledge of results versus work with knowledge of results. *Psychol. Monog.*, 1920, No. 125.
4. ELWELL, J. L., & GRINDLEY, G. C. The effect of knowledge of results on learning and performance. *Brit. J. Psychol.*, 1938, **29**, 39-53.
5. FRICK, F. C. Some perceptual problems from the point of view of information theory. In *Current Trends in Information Theory*. Pittsburgh: Univ. Pittsburgh Press, 1953.
6. GARNER, W. R., & HAKE, H. W. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, **58**, 446-459.
7. GELDARD, F. A. Adventures in tactile literacy. *Amer. Psychol.*, 1957, **12**, 115-124.
8. HAWKES, G. R. Cutaneous Communication; Absolute identification of electrical intensity level. *J. of Psychol.*, in press.
9. HAWKES, G. R., & WARM, J. S. Communication by electrical stimulation of the skin: II. The stimulus intensity range. *USA Med. Res. Lab. Rep. No. 401*, Fort Knox, Ky., 1959.
10. MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 1956, **63**, 81-97.
11. NEWBROUGH, J. R. Interaction between total stimulus information and specific stimulus information in visual recognition. *J. Exp. Psychol.*, 1954, **55**, 297-301.
12. POLLOCK, I. Message uncertainty and message reception. *J. Acoust. Soc. Amer.*, 1959, **31**, 1500-1508.
13. SHANNON, C. E., & WEAVER, W. *The Mathematical Theory of Communication*. Urbana: Univ. Illinois Press, 1949.
14. STEVENS, S. S. The direct estimation of sensory magnitudes: Loudness. *Amer. J. Psychol.*, 1956, **69**, 1-25.

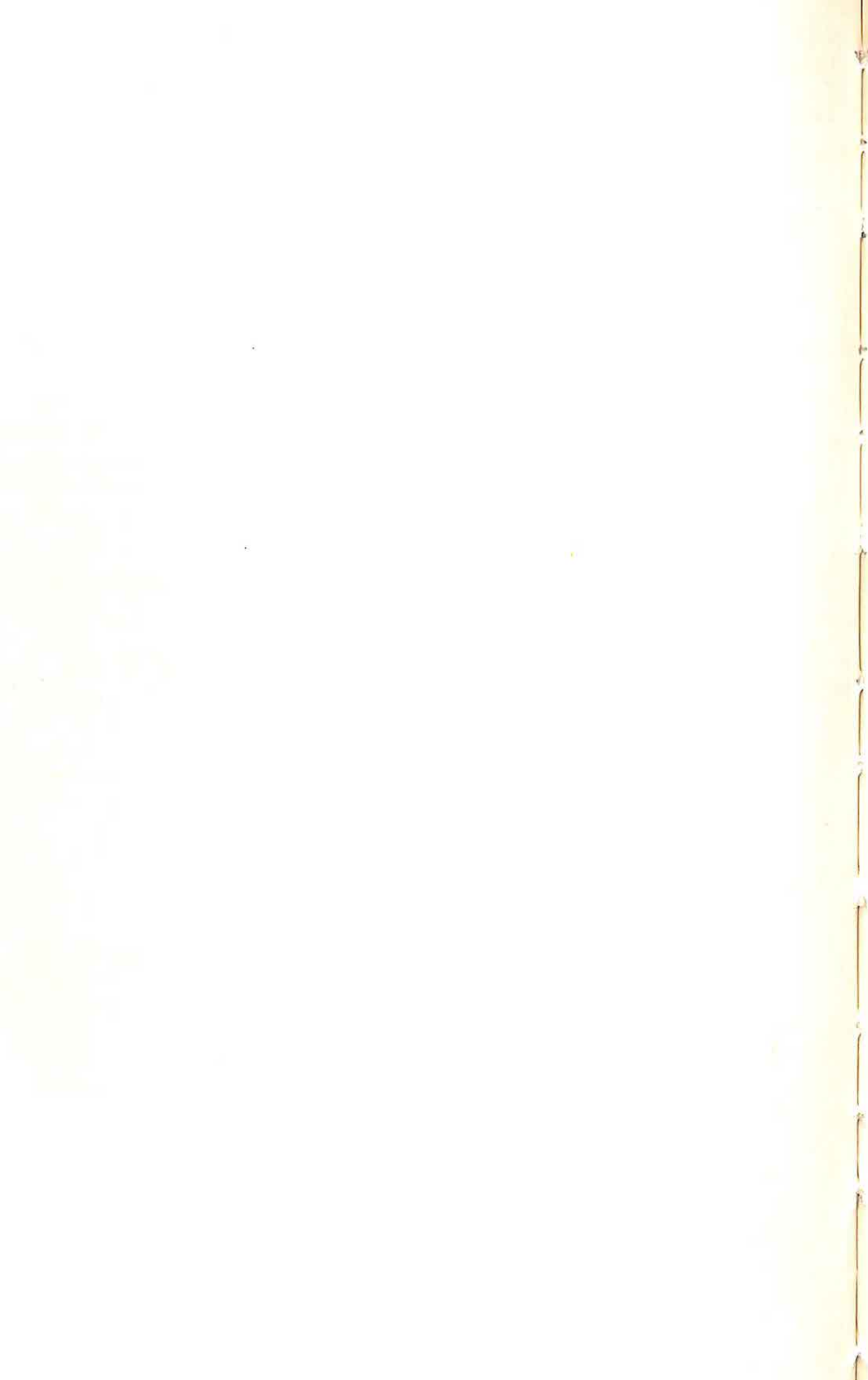
15. _____. Cross-modality validation of subjective scales for loudness, vibration, and electric shock. *J. Exp. Psychol.*, 1959, 57, 201-209.

Psychology Division

U.S. Army Medical Research Laboratory

Fort Knox, Kentucky





HEAD-BALANCE AND SITTING POSTURE: AN X-RAY ANALYSIS*

Institute for Applied Experimental Psychology and School of Dental Medicine, Tufts University

FRANK PIERCE JONES AND PHILIP F. M. GILLEY, JR.¹

A. INTRODUCTION AND PROBLEM

Jones *et al.* (3) described two erect sitting postures, an "habitual" and an "experimental," which can be defined operationally in terms of the way the head is balanced while the posture is being assumed. The two postures are illustrated in Figure 1, *B* and *C*. In both of them, the subject has moved up against gravity from the "relaxed" posture of *A*. *B* represents the

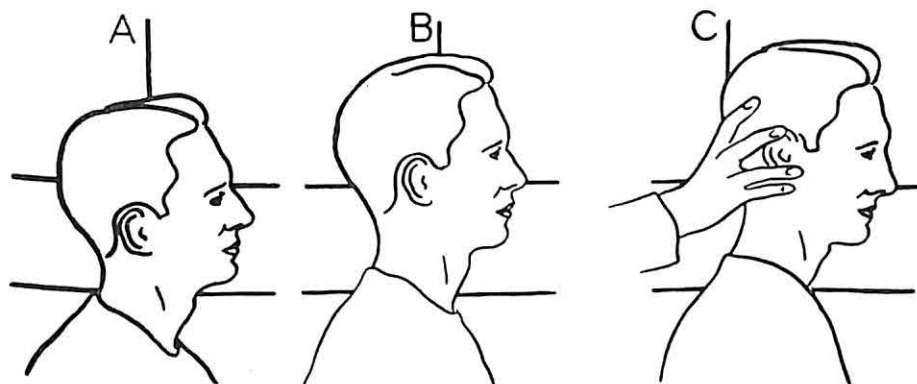


FIGURE 1
THREE SITTING POSTURES (TRACED FROM PHOTOGRAPHS)
A, Habitual relaxed. *B*, Habitual erect. *C*, Experimental.

subject's habitual response to the stimulus to sit up. In *C*, the experimenter has brought the subject's head into a new equilibrium while the movement was being made. Relative to the other two, the experimental posture may be described as "forward and up."

The operations performed in inducing the new, experimental posture are primarily inhibitory. The experimenter has prevented an automatic postural

* Received in the Editorial Office on February 4, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ This investigation was supported by grants from the Carnegie Corporation of New York and the USPHS, (RG-4836). Publication costs were defrayed by the Tufts Fund for Research in Kinesthesia.

response—a change in the reflex balance of the head—which appears in the average subject when he moves against gravity. It is a pattern of response which the subject can be trained to detect and inhibit by himself.

To the average subject, the experimental posture (*C*) feels different from that of *B* or *A*. There is a decrease (almost a disappearance) in the feeling of weight which is frequently associated with *A* and none of the effort which *B* ordinarily entails. This kinesthetic effect of lightness carries over into a subsequent movement, the pattern of which is correspondingly altered (3).

B. SUBJECTS

We have used x-ray photography to study the differences in the head-neck relation between the habitual and the experimental postures. Our subjects were 20 male students at the Tufts University School of Dental Medicine. They understood that the purpose of the study was to obtain data on the relation between posture and occlusion.

C. PROCEDURE

The subject was seated in a backless, dental chair of adjustable height. Three sagittal plane x-rays were taken while the subject sat in (*A*) his "most relaxed" posture; (*B*) his "most erect" posture; and (*C*) the "experimental" posture defined above.

To analyze the photographs, three planes were drawn on tracing paper clipped over the film. They are: (*a*) the Horizontal Plane—a horizontal line drawn tangent to the most superior surface of the spine of the axis (the second cervical vertebra); (*b*) the Bolton Plane—a line drawn through the nasion and the Bolton Point (the highest point in the depression posterior to the foramen magnum at the base of the skull); (*c*) the Vertebral Plane—a straight line drawn through the lowest point of the second and fourth cervical vertebrae in the line of radiopaqueness defining the distal ends of the vertebral foramina.

With these three planes for reference, the following linear and angular measures were made: (*a*) The Neck Angle—the angle made by the intersection of the vertebral plane with the horizontal. (*b*) The Head Angle—the angle made by the intersection of the Bolton plane with the horizontal. (*c*) The Head-Neck Angle—the angle made by the intersection of the Bolton and the vertebral planes. (*d*) The Height of Sella Turcica—the perpendicular distance from the center of sella turcica to the horizontal plane. (Sella turcica is defined as "the depression in the superior surface of the

body of the sphenoid bone which houses the hypophysis cerebri (pituitary gland)). It is said to bear a close relation to the center of gravity of the head (1, p. 191). (e) The "Post Spinal Distance"—the perpendicular distance of the most distal point on the atlas (the first cervical vertebra) from the horizontal plane.

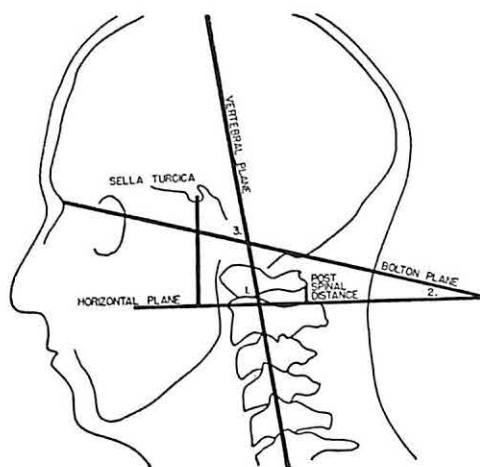


FIGURE 2

PLANES, LANDMARKS, AND ANGLES USED IN ANALYZING X-RAY PHOTOGRAPHS
1, Neck angle. 2, Head angle. 3, Head-neck angle.

The measures were taken for each of the three postures. Means and significance levels of the differences are given in Table 1.

TABLE 1
MEASURES USED TO DISTINGUISH X-RAY PHOTOGRAPHS OF THREE SITTING POSTURES

Index	Means			Significance of differences		
	Posture A	Posture B	Posture C	A vs. B	A vs. C	B vs. C
1. Neck angle	72.51	80.19	78.50	.001	.001	—
2. Head angle	16.74	24.30	15.61	.001	—	.001
3. Head-neck angle	56.71	55.78	62.46	—	.001	.001
4. Height of sella turcica	66.87	76.61	67.76	.01	—	.001
5. Post spinal distance	11.20	11.24	13.24	—	.001	.01

D. DISCUSSION

Index 1 (the Neck Angle) does not distinguish between the two erect postures. In both of them, the subject has straightened up against gravity and the angle between neck and horizon has increased. The other four indices, however, distinguish them sharply. In the experimental posture, the

angle between head and horizon is smaller, the angle between head and neck is larger, sella turcica (and by inference the center of gravity of the head) is lower, and the distance between the spines of the first two vertebrae has increased. In the light of subjective reports that the experimental posture is easier to maintain than the habitual, these measured differences (which statistically have a high degree of significance) suggest that a change has taken place in the balance of forces which control posture. To explain the change, force diagrams have been constructed and the following hypothesis is advanced:

The head ordinarily rests in unstable equilibrium with the center of gravity forward of the base of support. The torque thus placed upon the

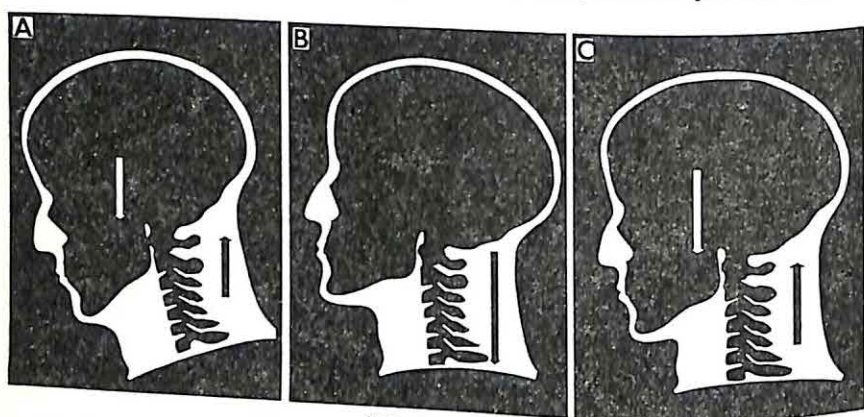


FIGURE 3
VECTOR ANALYSIS OF FORCES ASSUMED TO BE ACTING ON THE HEAD IN THREE SITTING POSTURES
A, Habitual relaxed. *B*, Habitual erect. *C*, Experimental. (Diagrams traced from x-ray photographs).

head by gravity is transmitted to the complex of muscles and ligaments at the back of the neck and through them to the spinal column. So long as the gravitational torque is allowed to operate, the head serves as a mechanism for maintaining tonus in extensor muscles and facilitating movements made against gravity. In the habitual erect posture, the head is rotated backward and the center of gravity lifted. This rotary movement eliminates the gravitational torque together with the stretch which it places on extensor muscles. Thus the erect posture is assumed and maintained without the mechanical advantage of tonic facilitation from the head. In the experimental posture, this advantage is preserved by the experimenter. Backward rotation of the head is prevented and the extensor muscles are allowed to lengthen during the movement. The well-known effect of stretch in increasing the lifting

strength of muscle (2, p. 122ff.) accounts for the feeling of decreased effort reported by subjects. (In this connection, it should be pointed out that this feeling is a function of the facilitated movement and not of the position as such. If, as sometimes happens, the subject straightens up in his habitual way and then changes the axis of his head into that of the experimental posture, he will increase rather than decrease the strain and effort which the habitual posture ordinarily evokes.)

Force diagrams are shown in Figure 3. Vectorial arrows indicate moments of force about the balance point of the head. In *A*, the head is in dynamic equilibrium, with the gravitational pull balanced by the elastic force of muscles and ligaments in the neck. In *B*, this balance has been replaced by a downward force, represented by a single arrow of double length. In *C*, the balance of forces has been restored, the increased length of the vectors indicating the lengthening of muscles which took place when the posture was assumed.

Further studies of the kinesthetic phenomenon are being made at the Institute for Applied Experimental Psychology and will be reported later.

E. SUMMARY

X-ray photography was used to study differences between two erect sitting postures which are sharply distinguished kinesthetically. The two postures were found to differ significantly in four indices: the angle of head to horizon; the angle of head to neck; the relative height of *sella turcica*; the distance between the spines of the first two vertebrae. An hypothesis was advanced to explain the kinesthetic phenomenon in terms of the balance of forces acting on the head.

REFERENCES

1. DEMPSTER, W. T. Space Requirements of the Seated Operator. W.A.D.C. Technical Report 55-159. Wright Patterson Air Force Base, Ohio, July, 1955.
2. FULTON, J. F. A Textbook of Physiology. Philadelphia and London: Saunders, 1949.
3. JONES, F. P., GRAY, F. E., HANSON, J. A., & O'CONNELL, D. N. An experimental study of the effect of head balance on patterns of posture and movement in man. *J. of Psychol.*, 1959, 47, 247-258.

<i>Institute for Applied Experimental Psychology</i>	<i>School of Dental Medicine</i>
<i>Tufts University</i>	<i>Tufts University</i>
<i>Medford, Massachusetts</i>	<i>Boston, Massachusetts</i>

PROBLEMS IN THE QUANTITATIVE EVALUATION OF PATIENTS WITH PSYCHOSES OF THE SENIUM*†

Rockland State Hospital, Orangeburg, New York

KENNETH LIFSHITZ

A. INTRODUCTION

In the New York State Hospital system 40 per cent of all first admissions are accounted for by patients with the diagnosis of either psychosis with cerebral arteriosclerosis or senile psychosis. Together these two diseases constitute what is commonly known as the psychoses of the senium. In 1936 this diagnostic group made up 28 per cent of all first admissions (4). These figures indicate the increasing importance of this group of diseases.

Despite the magnitude of the number of persons suffering from these illnesses there has been proportionately little investigatory and therapeutic research done with this group. There is a multiple of reasons why this is so, among which are: the feeling that this group of diseases is caused by an irreversible deteriorating process, the lack of knowledge of the nature of the processes involved, and the inherent difficulties of dealing with a deteriorated population such as this. However, there is indication that the conduction of research with this population is expanding. This has occurred as increasing medical knowledge has indicated that certain disease processes thought to be untreatable, such as atherosclerosis, which contribute to the pathology of the psychosis of the senium, are now yielding to therapeutic measures. It is also to be noted, in view of the fact that research has to be paid for, that the increasing socio-economic importance of the psychoses of the senium will hopefully make available increasing funds for the conduction of research in this area.

B. DISCUSSION

Among the problems involved in investigating patients ill with the psychoses of the senium the first two to arise are: (a) how to determine who has the disease and (b) how to determine the severity of the disease. The first problem is not one which permits of a definitive answer (1, 2) even at the autopsy table (8). In the usual case the differential diagnosis between the two psychoses of the senium is made by a hospital psychiatrist, who is

* Received in the Editorial Office on February 12, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

† This work has been supported in part by a grant from The Albert and Mary Lasker Foundation.

probably considerably overburdened with patients, many of whom are more interesting therapeutically than the senile population. Frequently no adequate history can be obtained because the patient cannot give one and there are no close relatives. If the patient is over 60 years of age, shows evidence of a psychosis, and has no history of previous major psychiatric difficulty, then a diagnosis of cerebral arteriosclerosis with psychosis is made if there is evidence of cerebral vascular accidents or other systemic arterial disease and a diagnosis of senile psychosis is made if this evidence is lacking.

The generally held opinion is that both cerebral arteriosclerosis with psychosis and senile psychosis are organically determined diseases. The former being the pathological diagnosis when damage to the brain results from a primarily vascular process and the second being the diagnosis when the damage results from a primarily parenchymatous process.

However, undoubtedly there are also dynamic factors (3) involved, and some feel that these are of paramount importance. If two people suffer the loss of exactly equivalent amounts of brain matter but one was "well adjusted" emotionally and the other was of poor emotional adjustment it would seem reasonable to expect that the first person would be better able to continue to function more adequately whereas in the second person the loss might well be the final straw which prevents even minimal adequate functioning. It would therefore appear to be the conservative point of view to consider, as does Rothschild (7) that both dynamic and organic factors play a rôle in the clinical manifestations of these diseases. It is doubtful that psychological tests can separate organic and functional impairment any more successfully than the clinician can. For a review of the validity of psychological test in brain damage the reader may be interested in reading Yates' article (9). In view of the above mentioned difficulties it is probably advisable to consider patients who are diagnosed as having one of the psychoses of the senium as a somewhat heterogeneous group suffering from more or less organic brain damage, which is characterized by varying proportions of primarily vascular damage or primarily parenchymatous damage and who are influenced by greater or lesser amounts of dynamic difficulties. The foregoing statement may seem to be rather broad, however, it is probably no broader than the statement which could be made covering similar large groups of medical illnesses such as the schizophrenias.

The second problem, that of quantitating the degree of impairment is perhaps the more important one since it is only if we are able to quantitate that we will be able to tell whether or not therapeutic measures have been successful.

For those patients who are relatively mildly impaired and who are non-hospitalized there exists a great number of psychological tests which are directed at determining the degree of organic brain damage. Reviews of some of these tests have been published by Klebanoff *et al.* (6) and Inglis (5).

Hospitalized patients are frequently used in initial investigations since they are more generally available and it is possible to see them for repeated evaluations. In the testing of new therapeutic measures commonly the poorest risk hospitalized patients are selected; since the dangers of the therapy may be unclarified and its benefits dubious, the investigator feels most comfortable working with those patients with the poorest prognosis. For this particular category of patients the quantitative tests which are generally available are geared at too high a level. Many of these patients exist in a near vegetative state and large proportions would score at the extreme low end of available tests. These severely damaged patients have very limited attention spans and tests which consume in excess of 15 to 30 minutes cause them to become irritable, inattentive, and frequently overtly hostile. The same hostility is evoked when the patients are called upon to attempt to perform in manners which are far beyond their capabilities or, when they are approached by an examiner who is required by the test format to be rigid in his manner of questioning. Those tests which are simple and quick to administer are generally limited to a narrow range of functioning and though they may be statistically correlated with overall performance leave one somewhat fearful that he may be missing a change which is occurring in more global functioning.

C. PRESENTATION OF A NEW TEST

At Rockland State Hospital we have been engaged in a study involving cerebral atherosclerotics. In conducting this study we have had to deal with the previously mentioned problems of evaluation. As indices of therapeutic success or failure we have been measuring death rates, general physical and cardiac status, incidence of cerebral vascular accidents, etc. We have also been faced with the problem of adequately measuring the psychological functioning of the patient. We consider the problem of psychological function of paramount importance since in the final analysis it is this impairment which has resulted in the patient's psychiatric hospitalization. In addition to other measures of psychological functioning we have been administering a test based on the time honored mental status evaluation. We have found this test to be simple, quick, non-taxing to the patient, reproducible over a time interval and, well-correlated with the other indices we have been using. Table 1 reproduces the test, which we call the Mental Status Check List, in its entirety. It

TABLE 1
MENTAL STATUS CHECK LIST

I. *Orientation*

- Question A. What is your name? (If only first or last name is given the other portion is requested.)
 Question B. What is today's date? (If only portion of complete date is given, the other portion is requested.)
 Question C. What is the name of this place you are in? (If only a portion of the correct name is given, the rest of the name is requested.)

II. *Successive subtractions*

- Question A. How much is 100 minus 3? That's right, and how much is 3 from that? Correct! Keep on going, subtract 3 each time from what you get.
 Question B. Now do the same thing taking 7's from 100.

III. *General information*

- Question A. Who is the President of the United States?
 Question B. Who is the Vice President of the United States?
 Question C. Who is the Governor of (the appropriate state)?
 Question D. What city is the capitol of the United States?

IV. *Identification of objects*

- Question A. (Patient is shown a watch.) What is this called?
 Question B. (Patient is shown a pencil.) What is this called?
 Question C. (Patient is shown a ring.) What is this called?

V. *Abstraction*

- Question A. What does the proverb "A stitch in time saves nine" mean?
 Question B. What does the proverb "It's no use crying over spilt milk" mean?
 Question C. What does the proverb "The restless sleeper blames the couch" mean?

VI. *Writing performance*

- Question A. Would you please write your full name on this piece of paper?
 Question B. Write also the sentence, "The dog chased the cat."

will be noted that we have taken several of the more or less standard mental status questions and prescribed the method in which they will be asked. The individual questions do not test single functions but rather a complex interaction of various types of mental functioning. There is no time limit to any of the questions and the next question is asked when it becomes obvious that the patient either knows the answer or doesn't. The scoring was empirically decided on before the administration of the test to the patients. The particular numerical value for each question was selected in an attempt to give more spread of scores at the extreme low end of the test. Table 2 indicates the manner of scoring. If the patient scored the maximum on each question of the Mental Status Check List he would have a total of 50 points. The

TABLE 2
SCORING OF MENTAL STATUS CHECK LIST

I. *Orientation*

- Question A. Two points for the correct first name, two points for the correct second name. Total possible, four points.
- Question B. One point for the correct year, two points for the correct month, one point for the correct day. Total possible, four points.
- Question C. One point is scored if the patient gives some indication that he understood the question such as a reply "building, hospital, or room." Two points scored if the patient replies with a single portion of the appropriate proper name or if he uses a more specific statement such as "mental hospital." Three points are scored if essentially the proper name is given but some minor error is made. Four points are scored for the complete and proper name.

II. *Successive subtractions*

- Question A. One point is scored if the patient gives any evidence on his initial reply of any ability to properly subtract. Two points are scored if the subtraction is frequently in error, exceptionally slow with some errors and/or shows a strong tendency towards backtracking. Three points are scored if the subtraction is at a reasonable rate of speed and contains less than three or four errors. The maximum of four points is scored if the subtraction is at a reasonable rate of speed and contains no more than one error.
- Question B. Scored the same as above.

III. *General information*

- Question A. Appropriate identification of president scored three points.
- Question B. Appropriate identification of vice-president scored two points.
- Question C. Appropriate identification of governor scored two points.
- Question D. Washington scored two points.

IV. *Identification of objects*

- Question A. Appropriate identification of watch scored two points.
- Question B. Appropriate identification of pencil scored two points.
- Question C. Appropriate identification of ring scored two points.

V. *Abstraction*

- Question A. Zero is scored if the patient shows no indication of understanding and/or gives a reply such as "I don't know what that means." One point is scored if the patient gives some indication of understanding, and attempts to give a meaning to the proverb. Two points are scored if there is some indication of abstraction. Three points are scored if the patient gives a reasonable abstraction for the proverb.
- Question B. Same as above.
- Question C. Same as above.

VI. *Writing performance*

- Question A. The patient is scored zero if he is unable to write at all or merely makes a few meaningless scrawls. From one to three points are scored based on the examiner's judgment of the overall legibility and spelling.
- Question B. Same as above.

actual score is multiplied by 2 to give a value in terms of a maximum possible of 100 per cent. The author recognizes that the scoring and method of administration of this test is not as specifically or detailedly described as is the case with most psychological tests. This is mainly intentional since the test was not conceived of as a rigidly precise thing, but rather developed as an outgrowth of a desire to lend comparability to those portions of a standard psychiatric mental status examination which were less subjective in nature. It covers a range of types of mental functioning in a very brief manner. A considerable amount of leeway is left to the individual examiner to adapt to the particular situation that he is in, so that if for instance the patient is hard of hearing the examiner may shout or repeat a question, or if the patient possesses some degree of language difficulty the question may be rephrased in a manner which might be more understandable to him. If the patient is distracted by an internal or external stimulus the question may be put to him again. To what extent this leeway functions is left to the judgment of the examiner.

Though there is certainly a place for precisely and rigidly prescribed psychological tests, the author feels that there is also a place for tests which are more flexible. This is particularly the case when the subjects to be examined are not in themselves precise. The situation might be considered analogous to one in which an attempt is made to measure the height of persons to $1/64$ ths of an inch when quarters of an inch may be quite sufficient and more appropriate. If one is limited by time the briefness of the test would permit of the handling of a larger number of patients than would otherwise be possible.

In validating the Mental Status Check List it has been administered to 78 in-patients at the Rockland State Hospital who were diagnosed as either senile psychotics or cerebral arteriosclerotics with psychosis. The patients selected were between the ages of 55 years and 75 years with a mean age of 68 years. The mean time in the hospital was 40 months. Patients who were completely bedridden were excluded as were patients who appeared to be in imminent risk of dying, who were deaf, blind, or unable to speak the English language. At the same time that the Mental Status Check List was administered to these patients they were also given five of the Wechsler Adult Intelligence Scale subtests. The Mental Status Check List was administered by three persons and the WAIS was administered by two persons. The five WAIS subtests were: information, arithmetic, digit span, picture completion, and block design. These subtests were individually scored and then corrected to the age scaled scores as indicated in the WAIS manual.

For comparison purposes these five scores were then added to give one overall sum. From the method of preparation of the age scaled WAIS scores the universe mean would be 50, the mean sum of the five age scaled scores for our population was 21.8. The comparative performance of the patients on the five WAIS sub scales and on the Mental Status Check List can be seen in the scattergram, Figure 1. The Pearson Product Moment

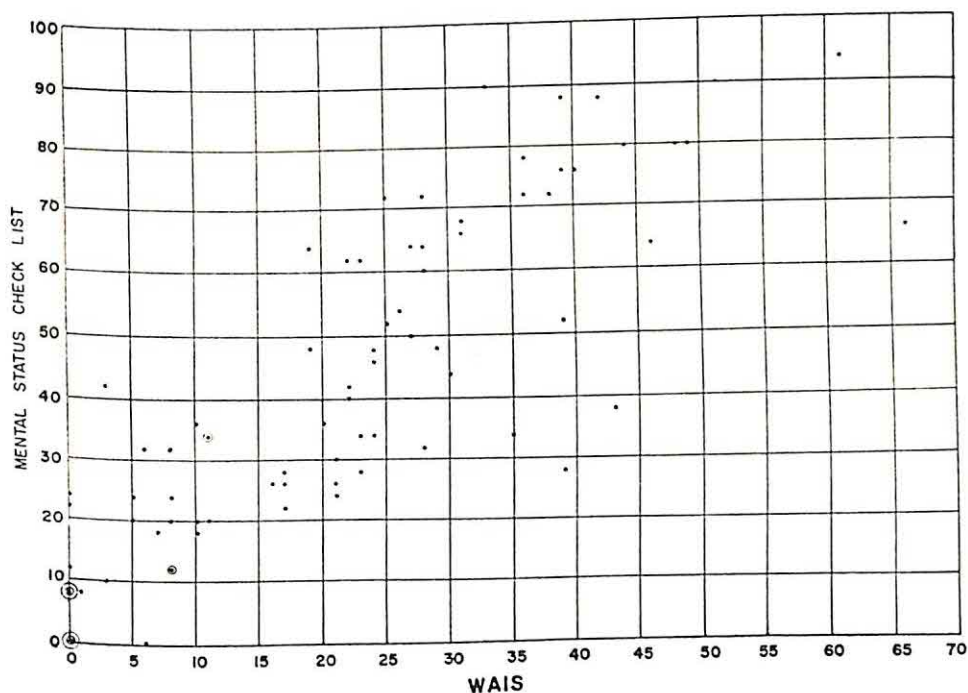


FIGURE 1

Scattergram of the scores on the MSCL (vertical axis) and the sum of the age scaled scores on the information, arithmetic, digit span, picture completion, and block design subtests of the WAIS (horizontal axis). Circles about dots indicate patients with identical scores.

Correlation between the Mental Status Check List scores and the WAIS scores was .82. The mean score on the Mental Status Check List was 41.36 with a standard deviation of 25.05. After a six month lapse of time 50 per cent of those patients who were still available for testing were randomly selected from the total group which was originally tested, and were again given the Mental Status Check List and the five subtests of the Wechsler Adult Intelligence Scale. The scores which these 26 patients received were then correlated with the initial scores. The correlation between the initial

WAIS testing and the WAIS testing after six months was .81 and the correlation between the two Mental Status Check List testings was .94.

The correlation scores achieved indicate that the Mental Status Check List correlates well with the sum of the age scaled scores of the five WAIS subtests used. The Mental Status Check List also correlates excellently with itself, after a six months time lapse, and to a greater extent than do the five WAIS subtests under the same circumstances. It therefore appears that it is possible for a relatively brief, simple, loosely structured test to yield valuable data with this type of population.

D. SUMMARY

The magnitude of the problem of patients ill with the psychoses of the senium is increasing. It is doubtful whether one can separate diagnostically the two illnesses, cerebral arteriosclerosis with psychosis and senile psychosis, which constitute the psychoses of the senium. Both organic and functional impairment play a rôle in this group of diseases. Of prime importance in evaluating the success of therapeutic measures in this population is a quantitative evaluation of psychic functioning. Many of the available psychological tests present problems in their administration to a severely deteriorated population such as the hospitalized senile psychotics. A test (the Mental Status Check List) based on a scored, partial mental status examination is presented. It has the advantage of being brief, 10-15 minutes for administration, and being structured in such a way as to test varied aspects of functioning in a manner which permits the examiner to adjust to the eccentricities of the patient. The correlation between the Mental Status Check List and the sum of the age scaled scores on five of the WAIS subtests was .82, with a n of 78. After a 6-months lapse of time the MSCL correlated with itself .94 and the five WAIS subtests gave a correlation with themselves of .81.

REFERENCES

1. ALTSCHULE, M. D. In *Cerebral Vascular Diseases* (Ed. Wright, I. S., & Luckey, E. H.). New York: Grune & Stratton, 1955. Pp. 48.
2. BARR, D. P. In *Cerebral Vascular Diseases* (Ed. Wright, I. S., & Luckey, E. H.). New York: Grune & Stratton, 1955. Pp. 72.
3. FERRARO, A. In *American Handbook of Psychiatry* (Ed. Arieti, S.). New York: Basic Books, 1959. (Vol. II, Pp. 1024.)
4. HOCH, P. H. (Commissioner) Sixty-eighth Annual Report of the Department of Mental Hygiene State of New York. Albany: Legislative Document No. 113, 1957.
5. INGLIS, J. Psychological investigation of cognitive deficit in elderly psychiatric patients. *Psychol. Bull.*, 1958, **54**, 197-214.

6. KLEBANOFF, S. G., SINGER, J. L., & WILENSKY, H. Psychological consequences of brain lesions and ablations. *Psychol. Bull.*, 1954, **51**, 1-41.
7. ROTHSCHILD, D. The rôle of the premorbid personality in arteriosclerotic psychoses. *Amer. J. Psychiat.*, 1944, **100**, 501-505.
8. SIMON, A., GOFMAN, J. W., MALAMUD, N., JONES, H. B., & LINDGREN, F. T. Lipoproteins in general and cerebral arteriosclerosis. *Amer. J. Psychiat.*, 1952, **108**, 663-668.
9. YATES, J. A. The validity of some psychological tests of brain damage. *Psychol. Bull.*, 1954, **51**, 359-379.

Rockland State Hospital
Orangeburg, New York

THE STIMULUS CONCEPT: *ANIMISM*?*

Department of Psychology, Michigan State University

LEO L. GLADIN

A. INTRODUCTION

The precise nature of the means by which the organism effects its contact with the objects and events of the physical universe has long been a matter of crucial importance for both philosophy and psychology. Within philosophy, the manner in which the organism becomes affiliated with its surround engenders epistemological questions; in addition, the systematic orientation of the individual philosopher depends at least in part on his viewpoint as to the degree in which the organism's sensory apparatus enables it to realize the external universe. Extremes of idealism or realism or something in between may be espoused, according to the philosopher's view of the matter.

Psychology, which espouses realism as a matter of course, must specify the avenue of contact of organism with environment for reasons of scientific necessity relating to the induction of *behavior*. Although epistemological problems may be deferred for later entertainment or relegated to the domain of philosophy, the problem of contact merits and is given early consideration in textbooks of general psychology. Within the first few pages, the "external stimulus" is specified as the means of contact; it is specified not only as the means of contact, but also as the instigator of response. Indeed, the stimulus is identified in terms of response, which leads to an often-criticized circularity of definition.

The external stimulus is variously described as "physical energy," "energy changes," or "anything"—always provided that it impinges upon and excites a receptor or receptor-system. The last description is more inclusive in that it takes account of such receptor-excitation as is not inevitably induced by means of physical energy or energetic change—as is the case with tactile sensitivity, in which the excitation may result from the exploratory activity of the respondent organism. More careful definition indicates that the organism plays some part in the *selection* of the external stimulus to which it will respond, as determined by its current drive status and structural organization.

Commonly, a distinction is made between *stimulus* and *stimulus-object*—

* Received in the Editorial Office on February 18, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

the latter being described as an aspect of the environment from which stimulus energies emanate or are reflected, thus to impinge on a receptor and excite it. Although the distinction appears logically necessary for purposes of accurate exposition, it is more often tacitly implied than it is explicitly used.

In addition to the above, introductory texts also make mention of *potential stimuli* which reside in the environment; these lack actuality either because they are not currently impinging their energies upon a receptor or because the hypothetical organism is not presently in a condition of need requisite to their selection.

A cursory comparison of general psychology texts will reveal that there is no commonly agreed-upon definition of the external stimulus, although this elementary concept is held to be the necessary link between organism and environment—the keystone of the arch of the behavioral sciences. There is a general agreement, however, that the external stimulus “puts something into” the organism which somehow provokes its responses.

Although the more sophisticated psychological literature does not often define the “input-output” function of external stimulation and observable response, the terminology employed strongly suggests that there is a well-nigh universal belief in the stimulative powers of the physical energies and other properties of the surround inhabited by the organism. *Stimuli* are said to “attract,” “distract,” “compete,” “arouse,” and “inform,” and to “produce,” “evoke,” “control,” and “elicit” responses. It will be noted that the predicate terms refer to animate and quasi-intentional functions of external stimulus-properties which proceed from the environment into the organism—so to instigate its responses.

It will be the thesis of the present paper to deny to the concept of external stimulation the powers traditionally ascribed to it, and to deny that such an event as an external stimulation of behavioral response can in fact occur. *The “external stimulus” will be labeled as a residue of animism which has been maintained within the behavioral sciences, and which has perpetuated a grave constant-error of conceptualization.* Initial discussion will be concerned with the evolution of the stimulus concept as it has been variously applied and unlawfully generalized, following which an attempt will be made to resolve the problem of how the organism makes effective contact with the physical surround.

B. CAUSE AND EFFECT

According to J. Müller: “Sensation consists in the sensorium’s receiving through the medium of the nerves, and as the result of the action of an

external cause, a knowledge of certain qualities or conditions, not of external bodies, but of the nerves of sense themselves" (2, p. 82).

The term "sensorium" is, of course, an archaic usage, and whatever objection may be raised concerning the doctrine of specific nerve energies which is implied in the statement, it nevertheless sets a limit within which the organism must operate, regardless of the nature of its contacts with the external world. For the purposes of the present discussion, the "qualities or conditions" . . . "of the nerves of sense themselves" may be viewed from the standpoint of central appreciation rather than qualitative specificity of the peripheral nerves themselves.

Under the above conditions, the organism seems to be effectively sealed off from its environment. However, the "action of an external cause" is granted. To an observer, an external cause which acts upon an organism produces an observable effect—a response. If an *effect* is a *response*, therefore the external *cause* is a *stimulus*. The argument is not logically sound, and probably has never been consciously advanced; nevertheless, an "appearance of things" which has not been seriously examined can be unconsciously maintained. The semantics of such phrases as the "attractive," or "noxious" stimulus reflect either a careless usage of language or an unconsciously maintained belief. An external cause cannot lawfully be treated as an external "stimulus."

The language of causal relations by which a cause is defined in terms of its effects also assigns to *cause* predicate values similar to those ascribed to the *stimulus*: it "produces," "impels," or "induces" effects. If now the observer interprets causal relations in terms of stimulus animations, he confounds his observational knowledge of cause-and-effect with something which is taking place within the organism, as if what he sees occurring must of necessity be a model of the stimulus function responsible for the subsequent behavior of the organism. He thus indulges in the "'psychologist's fallacy' *par excellence*" described by James in his criticism of observer-errors which unnecessarily (and inadvertently) complicate the accounts of psychology (4, p. 196 f.).

If a human experimental subject were confined to an area whose neutral gray floor contained a number of charged electrodes embedded randomly in its surface but invisible to the respondent organism, and if the subject were required to continue in constant locomotion, he would make responses "to" the causes represented by the electrodes. Ordinarily, the experimenter would allude to these as his "stimuli." Under the circumstances described, however, the charges of current are merely the *instrumental* causes of the subject's

leaping, etc., responses. Within the respondent organism, there resides a *motivational* cause: he is attempting to reduce or avoid painful sensations which his central receiving apparatus localizes in the soles of his feet. To be sure, a very real *physiological* stimulation of nerve-endings of pain sensation does occur, but this aspect must be ignored for the moment in the interest of clarity.

Rendering the electrodes "visible" by representing them as red discs, the experimenter will note that his subject now becomes quite successful in avoiding shock experiences, this new behavior pattern being caused by what the observer calls "visual stimuli," meaning physical energy of a type which stimulates receptors for vision and somehow "elicits" or "produces" the avoidance behavior of the subject. In reality, when the subject steps on a now visible red disc he experiences the familiar knowledge of pain, as a quality "not of external bodies, but of the nerves of sense themselves." In addition, by means of his receptors for vision, he experiences a discriminable quality of color-form which seems to "adhere" faithfully to the cause of his pain. He then learns to ascribe to the apprehended object, the red disc, the animate quality of his pain experience which continues to reside within himself. Having furnished the red discs with the animus of his pain, he is enabled to avoid the objects as he locomotes. It will be noted that the respondent organism ascribes to external causes a stimulus-value, and acts as if his subsequent responses were evoked by the external object.

The extensive employment of electrical current as a laboratory instrument itself provides a source of confusion as to the stimulation process; such current is an energy form, and when applied it looks very much as if the active input bears a necessary relationship with the subsequent response. However, the experimenter can substitute tiny needle points for his electrodes and obtain a similar consequence, even when his "stimuli" are inactive and depend for their function upon the operations of the respondent. If such instrumental causes are then made visible, not of themselves but by the addition of an aspect which is discriminable by the subject, his avoidance behavior will parallel that which was observed within the previous design. It will be noted that the respondent assigns an animus to a quality which seems to "adhere" faithfully to the instrumental cause of pain. Yet the animus which is assigned to "red-disc quality" can only be the instrumental cause for avoidance behavior, since it can yield no information as to the actual status of the instrumental cause of pain; the current may be disconnected, or the needle points withdrawn from the floor surface, but the only way in which the subject can discover this fact is by stepping on the red discs and perhaps

repeating the previous painful experience. Similarly, the antelope assigns to the *appearance* of the lion the animus actually resident in its alimentary predispositions; such antelopes as tended to validate the relationship of appearance to predisposition were, of course, less likely to survive and reproduce their kind.

The "psychologist's fallacy"—which seems to suggest that the behavioral scientist is somehow more naïve than his physicalistic fellows—actually represents an occupational hazard: his chosen field of study is the organism, yet he is himself an organism among organisms. His habitual manner of reacting to the objects and events of the environment may the more readily seduce him into observational error.

C. THE "STIMULUS" AS AN INSTRUMENT

In the illustration cited above, the charged electrode has a double significance: it is the "external stimulus," and hence an integral part of modern behavioral theory, and it is an *instrument* employed by an investigator who has good reason to believe that it will yield a specific consequence when it is appropriately applied. The latter significance is often overlooked by both reader and writer when the investigator reports his results and designates his "stimuli."

When introspective psychology was in vogue, the stimulus was regarded almost solely as an instrument; the experimental subject was trained to attend upon and report his "conscious sensations." As American psychology deserted the subjective in favor of the modern objective phase, Titchener remained as the lonely guardian of introspectionism. He saw his fellows devoting their attentions to external, physicalistic "stimuli" and coined the term "stimulus error" in protest. According to Boring: "Titchener's view was essentially that sensations are immediately given and that the nature of objects is inferred from them. Hence, to report on the stimulus itself is to report on an inference and to be giving interpretations . . ." (2, p. 610). The new "objective" psychologists thought his objection to be esoteric and quasi-philosophical, the product of an ultra-conservative mind still engrossed in an out-dated methodology. (Boring offers an extensive coverage of the *stimulus-error* controversy and its ramifications in his *History*.)

Although it was offered within the context of the sterile introspective method, Titchener's warning was valid and in complete agreement with Muller's statement, as previously cited, which presumably continues to command respect within the behavioral sciences. However, the fatality which causes the organism to ascribe stimulus values to the objects and events of

the external surround—to ascribe to “external bodies” its “knowledge of certain qualities or conditions of the nerves of sense themselves”—also makes the notion of “stimulus-input” seem plausible. It is beautifully physicalistic as well—therefore appealing to the new “objective” psychology, which was bent on creating a physics of behavior. Thus, the “appearance-of-things” gained further reinforcement from science itself.

D. THE ANIMISTIC CONCEPTUALIZATION OF EXTERNAL ENERGY

The instances in which inputs of sheer physical energy cause organismic reactions belong largely within the domain of physiology. However, there is no clear-cut line of demarcation between this domain and that of psychology—which also studies physiological functions as they bear on the behavior of the organism. But when physiological functions are *miscalculated behavior*, a source of error is introduced into psychology which furnishes a manifold of reinforcement to the concept of the “external stimulus.”

Both disciplines use the terms “stimulus” and “response.” Ordinarily, the physiologist applies the terms to relatively uncomplicated laboratory operations; thus, he may use electrical stimulation to obtain a reaction from a muscle-nerve preparation which he has dissected out of the living organism. For him, the term *stimulus* may refer only to a laboratory instrument, and it may also be legitimately regarded as the *cause* of the reaction he desires to observe and record. He is investigating the reactions of a *part* of an organism which he can treat as an object complete in itself. Even when he investigates the functions of a specific receptor-system in an intact organism, he may lawfully apply the terms *stimulus* and *response* to the appropriate operations of his research—as long as he treats with a limited functional apparatus. His term *stimulus* bears no necessary theoretical implications, nor does it raise epistemological questions, since he is not interested in what the total organism “knows” or “means” when the *part* under investigation reacts.

When legitimate physiological concepts are literally applied to quasi-psychological operations and subsequently to the behavioral repertory of the organism in its environment, actual lawfulness vanishes and only a beguiling appearance of legitimacy remains. By extension, physiological reaction becomes response, and response *is* behavior. Working backward from physiological reaction as “behavior” leads to a discovery of an external “stimulus energy” which is responsible for the elicitation of an item of “behavior.” By means of unexamined inferences, that which is true of physiological operations gains an appearance of truth within strictly psychological operations. Still, there is much that is attractive, from a physicalistic viewpoint, in the physiological

order of conceptualization. The reactions of *parts* are admirably invariant and readily observable, as is the stimulus-input of highly respectable physical energy. Physiological stimulus-response operations appear mechanistic—a straight-line function in which output (reaction) directly follows input (external stimulation). One need only separate the stimulus and response sufficiently to make room for the peripheral nervous system and a schema for the study of behavior is established. Historically, psychology has learned that when the stimulus-response functions of the several parts of the organism are summed, the resultant whole has a lacuna in its middle. The clean invariance of performance elicited from the separate parts by the physiologist vanishes, being replaced by the rather less predictable behavior-of-the-total-organism.

Nevertheless, a more sophisticated psychology continues to conceptualize the organization of behavior in a similar input-output fashion. The external stimulus is described as the Independent Manipulable Variable and the response as the Dependent Variable; between the pair of observable variables, the covert functions of the central nervous system are represented as Hypothetical Constructs or Intervening Variables. The directional sequence upon which behavior depends is thus given as:

—→S (IMV) —→HC (or IV) —→R (DV) = BEHAVIOR

The animation of the organism which is designated as *behavior* therefore continues to be dependent upon the animating energies and other properties of the external surround. In reality, only the *cause* may be regarded as external; when an observer imputes to that cause a contribution actually made by the organism—an animus which can be no portion of the indifferent physical energy which serves to excite the receptor apparatus—the schema breaks down. However adequate it may seem to be for purposes of experimentation, the formulation must inevitably nullify the products of experiment when it is applied to theories of behavior. In addition, the above schema creates a semantic confusion as to the actual function of physical energy when it is tacitly accepted by the behavioral scientist as he goes about his everyday professional affairs.

Physical energy is often active, and can *activate* the electro-chemical stimulation of neural processes. When the physiologist uses physical energy to “stimulate” a muscle-nerve preparation, thus animating his tissue-section with what he designates as a “response” there is given no warrant that this process is a model for the elicitation of such responses as are called “behavior.” When he activates a receptor-neural system and thereby elicits what he refers

to as a "sensation" there can be no warrant for the assumption that the process is a sample of "behavior" or that the process leads more or less directly to behavioral response by means of "... the automatic switchboard activities of the central nervous system" described by Hull (3, p. 40). When the physiologist applies varying intensities of "light energy" to the pupil of the eye and thereby causes the pupil to dilate or constrict, the resultant *adjustment of the part* cannot legitimately be regarded as an item of behavior. Yet when the behavioral scientist works with similar processes (as he does in conditioning experiments, for example) he unwittingly couches them in the semantic usages of his own profession under the tacit formula, "response-is-response-is-behavior."

From the standpoint of sound concept-formulation within the behavioral sciences specifically, the sensations, adjustments, and other reactions of the separately considered *parts* of the organism cannot be regarded as behavior. So regarded, such reactions help to maintain the fiction that behavioral responses are stimulated by the energies and other properties of the external surround. More than this, they help promote the fallacy of a conceptual schema which holds that behavioral responses are stimulated by means of *sheer physical energy*.

Relatively slight inputs of physical energy are said to "release," "increase," or "control" certain larger funds of physiological energy resident within the organism. Eminent behavioral scientists appear to hold at least implicitly with such a conceptual schema. The predicate terms quoted above are respectively those of Hull (3, p. 39), F. H. Allport (1, p. 409), and Lewin (5, p. 413). Only Lewin's intention remains somewhat obscure, since his animation of field-forces of the environment bears no absolute relationship with the more typical animation of physical energies *per se*. However, the terminology—with its semantic implications of underlying habits of belief—is widespread in the behavioral sciences.

Once the notion of *stimulus-input* is given credence, the function of *sheer physical energy* as the instigator of response becomes conceptualized, at least tacitly, in a univocal sense. By analogy, physical energy may be conceived as being represented by a straight stick which pushes a button (receptor-excitation) and thus sets off a chain of events which result in behavioral response.

A more accurate analogy would be one in which physical energy is depicted as an indifferently diffuse spray of a fluid complex. The fact of active motion of free energy is the *means* which enabled such receptor-systems as the visual and auditory apparatus to evolve at all, and which continues to enable

their respective functions. In addition, physical energy occurs in several ranges of frequency whose points of origin lie in the myriads of diverse aspects of the external universe. This feature enables the organism to evolve the central-nervous-system means of discriminating such aspects of the external surround as have significance for its well-being; if energy occurred in but a single wave-length, the organism could neither evolve nor utilize its complex discriminatory apparatus and would therefore have only an extremely limited capacity for the making of effective responses. Thus, while the "diffuse spray of a fluid complex" could not stimulate the behavioral responses of the organism, it can provide the *instrumental means* which indirectly enable the organism to sort out the several myriads of diverse aspects of its surround.

Nevertheless, more than one author has become enmeshed in the fallacy of *sheer* energetic stimulation of behavioral response—thus describing the energy reflected from words in print as resembling a small energetic charge which sets off a larger charge of physiological energy within the reader, evoking his response in this manner. And this in spite of his knowledge that a universe of difference lies between "eS?e" and "See?"

If Müller's dictum fairly approximates the true state of affairs in the animal kingdom, and the organism is indeed sealed off from the real world in the manner he describes, any external energy or other property of the physical universe having the actual power to "elicit," "produce," or "control" the behavioral responses of the organism must either have an intrinsic motive for doing so, or operating instructions from a benevolent and/or malevolent agency (or agencies) somehow motivated to establish contact with the strangely isolated organism. Thus far, physical science has discovered no supporting evidence either for agency or intrinsically motivated energetic forces, and in fact psychology has long since ceased to account for the varieties of behavior by such means.

Nevertheless, several generations of behavioral scientists have been trained to the notion that *In the beginning was the stimulus . . .* The habit of belief in the instigatory rôle of the "external stimulus" is so strongly ingrained that serious inquiry into its functions almost seems superfluous. It pervades schools of thought as diverse as Gestalt, Neo-perception, and Neo-behaviorism. Some of their representatives struggle against the notion without realizing precisely what it is that entraps them (as does Lewin by means of his Gestaltist field-theoretical approach which substitutes *force* for *energy*). Others appear to be quite comfortable with the concept of external stimulation of behavior. Everyday experience plus professional training combine to persuade them that the concept represents a real state of affairs rather than an "appear-

ance-of-things." Nonetheless, it remains the obligation of a science to penetrate illusions, however comfortable they may seem to be, and to discover *things as they are* out of *things as they seem*.

E. THE AHISTORICALLY CONCEPTUALIZED ORGANISM

The physical scientist can sum a number of molecules and obtain therefrom a predictable mole whose behavior is invariant. Even in theory, the behavioral scientist cannot sum a number of animalcules and predict therefrom the behavior of the resultant animal. Each organism is assembled as its evolutionary history dictates, and its behavior is determined by its particular experience of phylogenesis. The physical scientist bears no subjective relationship to the material under his study; unfortunately for the behavioral scientist, he does, and is required to engage in a continuous and vigilant struggle for disengagement. The material under his study is far more mutable and it may even alter its behavior *because* he is studying it. Where stimulus conceptualization is concerned, however, he may most readily come to grief when he persists in viewing the organism ahistorically—as it exists in its current state of integrity within the contemporary surround.

The psychologist may designate the red light of a traffic signal as the "stimulus" whose energy elicits the *stop* response of the automobilist. He may refer to the pistol-shot as the "stimulus" whose energy gets the runners in a foot-race into motion, and may even hint that the "stimulus energy" somehow becomes translated into the physiological energy of the running-response. At the same time, he knows that what he designates as "stimuli" are in fact signal-conventions, or "stimuli-by-appointment," designed by man for the purpose of regulating certain of his affairs. The energy itself continues indifferent and can have no vested interest in the *starts* and *stops* of mankind. By convention, an automobilist might be required to stop at any intersection which had no signal-light, and a foot-race could be started upon the cessation of the warbling of a nearby canary—under which circumstances the absence of an external stimulus energy would serve as the "stimulus" for the appropriate response.

The nativism-empiricism controversy—which began in a time when it was believed that each species of organism was created complete with all its structures and functions and thereafter remained immutable for all time—continues to the present day. That which is ahistorically conceived as being innate to the organism is referred to as "given" and only that which is conceived as having been learned during ontogenesis is referred to as "acquired." Since the "giver" has long been denied an explanatory rôle in

the accounts of psychology, the "given" has come to mean, "That which has not been acquired during the life-history of the organism." However, the "given" continues to impede accurate conceptualization and theoretical formulation. The more complex the organism, the more it seems to have been "given." The mammal used in laboratory experiments is ahistorically described as being prone to respond "naturally" to a "given" stimulus energy; meat powder "naturally" elicits the salivation response from the dog, and electric shock administered to a foot-pad "naturally" elicits the leg-raising response. Actually, the experimenter knows very well that the experimental subject must have acquired its response structures and functional capacities during phylogenesis and that the physical energies of the environment did not lose their indifference in the interest of furnishing guidance to the dog.

When Pavlov published his findings on laboratory conditioning processes, physiology carried its formerly legitimate terminology into a twilight-zone which is now claimed as the rightful province of psychology. The stimulus-response physiology properly applied to the reactions of the parts of the organism began to be applied to the responses of the total organism by the physiologist engaged in conditioning experimentation. The terminology was expanded to describe the newly discovered phenomena which were specific to the conditioning process, and was taken over intact by the young *objective psychology*. A misleading nomenclature thereby was introduced into the behavioral sciences and made its contribution to the concept of external stimulation of response by means of physical energy. Reinforcement was furnished to the notion that behavior is the product of a straight-line function—proceeding from the external world via quasi-observable *stimulus-inputs* and passing through the organism to return once more into the world as observable *response*.

The language of laboratory conditioning strongly suggests that the physical world has endowed certain of its energies with the power to elicit from an experimental animal "naturally" such responses as the investigator desires to cultivate artificially. However, when the subject organism—a dog, for example—is strapped into the harness which keeps it within the conditioning apparatus, it reacts at first to an electrical shock to a foot-pad by apparently attempting to escape the entire situation. This initial "natural" response is replaced by the natural response which is required in the experimental design and the animal learns to lift its insulted paw in response "to" what appears to be an input of the "natural"—or "unconditioned" stimulus-energy. This *appearance-of-things* seems entirely reasonable to an interested observer who

has accustomed himself to equate *external cause* with *stimulus*. Regarding the situation from the internal aspect of the dog, a neutral observer may see the animal's initial struggles to escape a situation which implies injury as entirely natural and the so-called natural response designated by the experimenter as somewhat artificial. When the dog's escape attempts are frustrated by its harness, it is impelled to react in a more discriminate manner within the limited situation. Experiencing sensations of pain, it localizes them appropriately and lifts the insulted paw under the stimulus of pain-sensation appreciated centrally *and as centrally represented*. During its phylogenetic history it has "learned" that threatened tissue-injury (such as is centrally represented as pain-sensation) is externally caused. That is, by separating a member whose nerve-endings are relaying their receptive-excitations to its psychological center, from a point of specific contact with a portion of the physical surround, the organism experiences a cessation of pain sensation. Its act of separation thus implies that some sort of operational knowledge of a property of the external surround has been gained during phylogenesis and/or ontogenesis.

Being paired with the "unconditioned stimulus" in temporal contiguity with the "unconditioned response" the selected "conditioning stimulus" begins to *elicit* a similar response. A previously neutral physical energy (buzzer sound, for example) has seemingly acquired an animus which can now act in a specific manner on the response apparatus of the dog. Of course, no reputable experimentalist would agree with such a statement. However, the semantic usage which places the *conditioning* within the organism and locates the *stimulation* outside of it may lead to unexamined unconscious inferences which are reinforced by the observer's witnessing of repeated instances of apparent input-output events.

The semantic dilemma is similar when the phenomenon of "stimulus generalization" is discussed. Obviously, the *generalization* takes place within the respondent organism, although the associated *stimuli* are designated as events originating in the external surround. The implication is that a physical energy which has previously become endowed with an animus to elicit a specific type of response also develops the power to share its animus with previously neutral energies. But this nonsense statement carries a graver implication: since by *response-occurrence* an event of stimulation preceding it is implied, the experimenter is moved to isolate the external cause and designate it with a stimulative animus. Thus, the dog which has been conditioned to raise one leg upon the presentation of buzzer-sound may also perform the conditioned response upon the accidental presentation of *door-*

slam should this sound occur within the experimental context. Since it has been conditioned to respond upon presentation of buzzer-sound, the dog may be said to have ascribed an animus to this sound; if it then responds upon presentation of door-slam, this does not signify that the accidentally introduced sound has been ascribed a similar stimulating animus *by the animal*. Yet the occurrence of the conditioned response leads the *experimenter* to assign the animus to the sound of door-slam, as a "stimulus-generalization" because he has noted the subsequent occurrence of the response. From the internal aspect of the dog, only the conditioned response has stability once it has been organized; the "stimulus," lacking an intrinsic animus, has no specificity and is therefore unstable. Any of a number of events of *physiological stimulation*, when centrally received within the experimental context, may set off the prepared response.

The phenomenon of "pseudo-conditioning" is not radically different although it represents a problem when the external causes are equated with stimulation. The ahistorically conceptualized "unconditioned response" is organized by repeated applications of "natural" or "unconditioned" stimulation—as when an infant is trained to withdraw its foot from contact with a charged electrode. When buzzer-sound, which has never been previously paired with shock, is then introduced by itself into the experimental context, the prepared response may occur, since it is the stable factor in the situation and labile to any number of centrally presented accidental events.

Ahistorically regarded, the phylogenetically prepared response designated as *startle reaction* looks very much like a reaction to an external stimulus; any number of external causes whose energies differ in kind and quality are said to "stimulate" it. Their common denominators are "suddenness" and/or "unexpectedness"—which are neither energies nor properties of the physical environment itself. To the mammal in its native habitat, the accidental event which is centrally represented as "sudden" or "unexpected" may or may not have profound significance for its future well-being. Since it cannot take time to assess the probable significance of such an event, it must come to full emergency alert under the built-in premise that danger threatens. At no set time, in no certain place, the external event whose progress remains undetected by the respondent until the point whereat actual physical encounter is possible retains a destructive potential. Out of a thousand such events, perhaps only one will destroy—but once is sufficient to a single organism; it must somehow learn to base its emergency reaction on a premise that any of the thousand events promises the fatal consequence. But the premise was "learned" during its phylogenesis; those organisms of its kind which failed

to animate their sudden or unexpected events with a probability of harm reproduced themselves less often, if at all, so that only the *quick* came to a current representation upon the field of reality.

A special characteristic of the startle reaction is that it may not recur upon a second presentation of the "stimulus"—although its quantity may be exactly equivalent to that which "elicited" the first reaction. The animus resident in energy which can instigate a violent reaction on one occasion loses its power on a second—a fact unknown to the physical sciences. Of course, it is equally unknown to the behavioral sciences and is only semantically implied on the basis of an ahistorical conceptualization of the stimulus function.

The modified form of alertness designated as "involuntary attention"—which is said to be "aroused," "attracted," and "provoked" by external stimuli—is not different in kind from startle reaction. The ahistorically conceptualized causes, *size*, *coloration*, *intensity*, *movement*, etc., can be envisioned at least as having some form of independent existence in the physical surround regarded apart from the receptor-neural systems of organisms; other causes, such as *novelty* and *repetition*, remain entirely dependent upon the temporal-spatial experiences of organisms. There is no *novel* physical energy, and repetition requires a witness to that which has gone before. Of the more palpable causes, it need only be observed that if the organism had not "learned" to react with interest upon being confronted by differing aspects of the physical surround during its phylogenesis, its kind would more likely be an object of study for the paleontologist than for the psychologist—save for some few human specimens which are protected within special institutions.

"Intensity of stimulation," which causes both repeated startle and persistent "distraction" of attention requires special consideration. It seems to offer proof of the external stimulation of behavior. However, the reactive structures which have enabled the organism to respond with sufficient effectiveness to continue in representation on the contemporary field also can present it with some material difficulties. Both repeated startle and persistent distraction as caused by intense physiological stimulation represent an overpowering of the sensitive neural apparatus which interferes with functional efficiency. Such stimulation may perhaps be best regarded as a *material cause* of mechanical reaction rather than as a behavioral "stimulus."

The arbitrary extension of the "life history" of the Organism, thus including the structural and functional acquisitions of countless generations of organisms into "Its" experiential repertory, is of course solely an expository

device; it does not imply a dismissal of the distinctions between phylogenetic and ontogenetic acquisition of the coping functions of organisms at large. However, the "extended organism" is a counteractive to another fictional conceptualization, the "ahistorical organism"—whose behavioral responses appear to be instigated from the outside world. Within this latter conceptual framework, the *stimulus error* can be negated, the external causes of behavioral effects can be plausibly animated, and the terminology and operations of the physiologist's laboratory can be made to appear applicable *en toto* to the domain of general psychology. While the total result looks physicalistic, an examination of any of the parts of this end-consequence reveals that the input-output schema is out of joint with everything the physical sciences have teased out of the appearance-of-things.

F. SOME SUGGESTIONS TOWARD AN AMENDMENT OF THE STIMULUS CONCEPT

Should all of the possible sources of behavioral stimulation be relocated within the organism—where, indeed, some of them are reported to be situated—the result would be more in harmony with the physicist's statements concerning the indifferent nature of physical energies and the other properties of the surround. In addition, the sources of behavioral stimulation would become relocated within convenient proximity to the mechanisms of response *qua* behavior—to which no one denies an internal locus. In satisfaction of the law of parsimony, the "potential stimuli" which bide their time in the environment, as if animistically awaiting their moment of actualization, would also become situated inside the organism. Once the animated "physical stimulus" is deprived of its power to induce behavior, the objective and subjective psychologies would be able to establish a more common meeting ground—no mean gain in itself.

Without the external stimulus, the "input-output" schematic order of behavioral organization collapses; all that remains is the external cause and the varieties of behavioral effect, which must now become dependent upon the organism itself for their elicitation. A more appropriate schematic order becomes necessary to account for such a status of affairs.

Were the sources of behavioral stimulation external to the organism, it would be necessary to postulate a number of spheres of influence—"worlds within worlds" of animistic agencies. The world of the nematode would thus be the special province of one agency, and the world of the sparrow that of another—thus to insure that the worm would not be stimulated into the behavioral reactions of the bird.

Within the tense equilibrium of the natural economy, all organisms of a single species of the animal kingdom do tend to behave in much the same manner; there is little time allowed them for the making of decisions. Once an emergency is over, they cannot wonder whether some other response might have been more appropriate than that which was made. A response is sufficiently appropriate if they survive to repeat it. Removed to the psychologist's laboratory, they can be induced to vary their responses according to an experimental design which is imposed upon them, and demonstrate the capacity to make responses requiring a finer discrimination than they would have time for within their native habitat—(of course they would also lack the presence of the laboratory apparatus there, as well as the necessity for using it). But, when they are in the experimenter's protection, their responses begin to vary within the species. A single "stimulus" may produce a variety of responses.

In the human's relatively protected surround, responses tend to vary even more, and often go beyond the bounds of "lawful" behavior as it is viewed from the standpoint of the natural economy; the human may court pain aberrantly, and such behavior would result in the swift extinction of a lesser animal and its kind. The net result is that psychologists do not often see *infrahumans* which love pain.

If the universe contained stimulating properties of a gross nature, it might be expected that all organisms would evolve similar behavioral proclivities, since the guiding agency would not then be regarded as a subdivided instigator of behavioral response according to species classification; therefore interspecies patterns of behavior ought to be found similar one to another. The lamb might be discovered practising its lion-hunting technique, or the lion might be discovered subsisting on grasses as does the lamb—and they might even lie down together. And all men might become savers of string.

It is obvious that the properties of the external universe lack this imperious and grossly indiscriminate domination over the several species of the animal kingdom.

The often-made qualification that each organism selects the external stimuli to which it responds according to its kind still begs the question; if the external energies and other properties are indifferent and neutral in the struggles of the animal kingdom, consistency admits of no qualifications. The properties of the surround must be regarded as an indifferent constant which cannot induce the behavioral responses of organisms.

However, the properties of this indifferent manifold are such that the organism is enabled to have access by their means to the external universe,

if it chances to evolve correlative receptor means. The indifferent manifold can neither deny nor offer; it is up to the organism's phylogenesis to devise the means of access, and the degree to which it is thereby enabled to cope with the objects and events of the surround remains dependent upon the discreteness of its structuring both on the affector and effector sides and within its psychological center.

The proposition which locates the behavioral stimulus external to the organism discourses on two levels at once. By invoking the physical energies and other properties which excite the receptor apparatus of the organism, it thereby describes the physically instrumental means which are correlative to the physiologically instrumental means by which the organism gains *access* to the objects and events of the external universe. But the proposition is not stated in terms of access; rather, allusion is made to responses of the organism elicited by external "energy," "energy changes," or "anything" which excites the receptor apparatus.

At a second, and more complex, level of discourse the "stimulus object" is brought under discussion, although not always under that designation and not usually until the general psychology text begins to refer to "perception." It is at this level of discourse that the *significance* which "things" have for organisms is invoked to account in part for the stimulus value of the objects and events which "stimulate" the responses of organisms. Although the two levels of discourse confound *access* and *significance* by a devious synthesis to yield a stimulus "complex," they remain incongruous if seriously examined; taken together, they bear a strong relationship to the ancient conceptualizations of external stimulation by means of the *simulacrum* and *ideolon*.

The two levels of discourse are further confounded when the behavioral "stimulus" adheres to the energy and is not referred to the *object* by the respondent organism. Hearing the roar of a lion, the antelope behaves as if it "knows" that the sound emanates from a lion. Hearing the ring of the telephone, the householder goes to the instrument from which the sound issues, but when he hears the door-bell ring he goes to the door and not to the bell. According to agreed-upon social conventions, signals are arranged which bear no necessary relationship to the emanating object. Should the householder try to "answer the 'phone" when he hears the door-bell ring, he will afterward acknowledge that he has made a "mistake."

In the conditioning laboratory, the dog which has learned to salivate upon the presentation of buzzer-sound does not refer the sound to a specific object, the buzzer from which it emanates; the experimenter has designed the situa-

tion so that the sound will become associated with meat-powder, whose appearance is the "natural" stimulus for salivation. Apparently because meat-powder seems to be specifically related to the energy which makes it visible to the dog, the experimenter assigns a similar specificity to the energy of buzzer-sound; when the animal responds "to" "accidental-door-slam" with salivation, he does not infer that the physiological apparatus of the respondent has somehow made a mistake but describes the event as a "stimulus-generalization" which spreads the specificity over a range of external occurrences. On this basis, the question of "mistake" vs. "generalization" still continues to be discussed (6).

In point of fact, the energies which render the environment accessible to the visual, etc., receptor-systems of the organism cannot guarantee the validity of external objects and events; if they could, they would be *eidola*, or faithful copies of objects and events. If the dog makes an "unconditioned" response "to" the appearance of meat-powder, it can only happen because in its past experience certain appreciated qualities have been more or less faithfully represented as the properties of external objects and events. It may be supposed that, within the experimental context, the dog which has learned to salivate upon the presentation of meat-powder will also afterward salivate upon the presentation of a small block of wood until its combined alimentary and visual experiences enable it to discriminate between the two. Sheffield's studies of the effects of saccharine on learning processes (7, 8) indicate that the phylogenetic acquisitions by which "sweet-taste-quality" represents a nutritional characteristic regularly attendant upon certain objects of the external environment, serve to beguile the organism into responses which would lead to its extinction within the native habitat (if it could elect to attempt an existence by the guidance of receptors for taste alone).

The psychology of perception, which has made much progress in the exposition of the rôle of *significance* of objects and events as the determiners of behavioral responses, is nevertheless encumbered by the traditional stimulus concept. Its literature is sprinkled with references to *inputs*, *energies*, *information*, and other attributes of the "external stimulus." Although its experimental products have made a definite contribution to psychology's fund of knowledge, it nevertheless is hampered in much the same fashion as is behaviorism in its efforts to draw the threads of research into a plausible fabric—and for much the same reasons. The functions of *significance* are contaminated by their ultimate association with the functions of *access* in the traditional rôle of "external stimuli." Therefore, the following presentation, which has much in common with the psychology of perception, must exclude it from consideration at least on semantic grounds.

G. THE PROCESS OF STIMULUS ASCRIPTION

If the energies and other properties of the external surround can be neither grossly nor selectively the animators of behavioral response, accounting must be made for the *appearance-of-things* which pervades not only the language of the laboratory but also the common form of speech—whose usages are frankly regarded as animistic by the behavioral scientist. Such phrases as “the sea beckons,” and “those peanuts are *so* inviting,” reflect more than a poetic turn of speech; underlying them is a near-universal tendency to attribute to the object a function of the subject which is the stimulative of his behavioral responses. This transposition is an articulation of the mechanical function which enables the respondent to address his reactions toward the objects and events of the spatially extended world rather than to the “nerves of sense themselves,” since this solipsistically limited mode of reaction can yield to the organism no advantage from the world and must inevitably result in its extinction by some object or event of the world.

By means of its implicit transpositions, the organism charges the objects and events of its environment with various degrees of significance, according to a hierarchy of values somehow acquired by the respondent during its phylogenesis and ontogenesis. An intact, well-integrated human (by way of illustration) may take a notion to explore some strange terrain; within his visual field he will discover certain portions of his own somatic periphery already familiar to himself, plus a grassy plain, trees, hills, and a sky. Although he has never experienced this specific terrain before, it bears a general similarity to landscapes of his past experience. His organismic capacity to “generalize” reduces the quality of foreignness of his present experience, although it does not alter the absolute fact that he has never experienced this specific terrain before. His capacity to generalize cannot be regarded in terms of “mistake” as previously was discussed, since his organism has somehow learned to classify certain aspects of objects and events according to such qualitatively appreciated degrees of similarity as seem to “adhere” consistently to actual objects and events wherever and whenever he may chance to experience them.

The landscape within his visual field may be structured by the respondent as a neutral background which can do nothing *for*, *to*, or *with* him. There need be no reason for him to respond “to” it with items of behavior, even though many of his receptor systems are active and he is being literally bombarded with the energies of the environment. The myriad odors, ranges of “light” and “sound” frequency and *et cetera* which impinge on his receptor-periphery seem unable to elicit his behavioral responses. Since he is a complex

and sophisticated organism, most of their effects have been neutralized via *adaptation*, either peripherally or centrally.

Should the energetic complex which represents a *snake* in his visual field become reflected on his retinal apparatus, even this energetic complex seems unable to elicit his response until he "notices" it. However, by noticing the snake he does not *select* a discrete energetic complex and thereby render its "stimulating animus" operative. If he has never had experience of any kind whatever relative to snakes—a highly improbable circumstance—he may be regarded as being in the presence of a "novel stimulus" which "attracts his attention." Actually, it is a specific lacuna in his experiential repertory which could be held to better account as the stimulus for the focus of his attention on the snake, since the "novel" has an unknown potential whose value must somehow be assessed when it becomes intruded into the ongoing transactions of organism and environment. If now the snake activates certain buttons attached to its tail, the human may attribute the rattling sound to an unseen grasshopper whose "noise" is familiar to him, since he cannot be expected to recognize the attributes of a phenomenon he has in no manner previously experienced.

Usually, the human has had some sort of experience with snakes, either verbal-affective or actual; the nature of that experience results in the *ascription* of a *stimulus value* to the energetic complex which represents the snake within his visual field. The kind of response which is elicited is dependent not upon the "energetic input" but upon the nature of the stimulus ascription, and the subsequent reaction may range from precipitate-flight-in-terror to picking-up-a-stick-and-bold-advance.

The respondent who has had no snake-experience may approach and attempt to handle the object of his interest and thereby be injected with venom. Should he chance to recover from this event, he will thereafter attribute to a specific complex of color, form, movement, and sound which attended upon event-of-snake-bite a *potential value* associated with his experience of pain-and-suffering. This value will become the *potential stimulus* which will accompany him on future expeditions into terrains resembling the locus of his actual experience; however, he will be on guard as if the "potential stimulus" really has its existence in the environment. For him, *sticks, movements-in-the-grass, and noises* may well develop a stimulating animus which effectively solicits his avoidant behavior. He has in fact *reversed* certain qualitative accompaniments of a significant experience of the world upon the world of his future perambulations.

H. THE PROCESS OF REVERSAL: FROM SUBJECT TO OBJECT

The human infant discovers the boundaries of its somatic periphery by a series of largely accidental explorations, touching itself here and there and noting the effects of its contacts upon itself. It also takes notice of contacts made upon it by the external surround which feeds, bathes, and otherwise keeps it tidy, so that some of the effects it experiences are not self-induced, as it comes to learn—which fact no doubt aids the infant in its establishment of an implicit faith in the existence of an external world beyond its periphery.

Despite some sophistication, when the infant begins to explore the actual surround it is stimulated by the vast lacuna in its experiential repertory; almost the whole of the environment is novel to its experience. It embarks on its explorations without much design; its behavior is largely accidental, so that it may be said to experience at first nothing more than a series of collisions with the environment. Such collisions cause it to experience much pain and much pleasure, and if it is a normally endowed infant, its behavior becomes more patterned and less accidental; of the two alternatives, it usually seems to lean toward the pursuit of pleasurable experience and the avoidance of that which occasions pain. From its sensating core, it begins to endow the objects and events of its experience with stimulus values, according to the significance they have *for it*. The direction of organization proceeds *from* the experiencing subject *toward* the experienced objects according to more and more discriminate classifications of the centrally appreciated qualities which attend its experience of external objects and events. As the various aspects of the objects and events come to be discretely defined by means of the access furnished through its several receptor-systems and the inevitable accompaniment of a gamut of responses, the developing infant (without giving any thought to the process) begins to endow the indifferent manifold of physical energies and other properties of the surround with specific animation, and begins to act as if it were somehow stimulated to respond by external objects and events. Through its experiences of sensation and response, it has reversed the original sequence of events and has animated the world with its own sensations.

I. THE WORLD "INSIDE"

Since only infants and the victims of severe neuropathology experience the world as sensationally resident within themselves, not much thought is given to the fact that the world *as sensed* is internal to the organism. Whereas response operations continually confirm the organism's faith in the

existence of the external world, its receptor-neural functions—which can proceed only in a centripetal direction to a terminus within the central nervous system—as inevitably locate the world-as-sensated within the organism. When regarded from the standpoint of effective response-behavior, the schema is an excellent one, since the neuro-physiological representation of a present experience is always in potential contact with the vast repertory of past experiences, in whatever way they may be represented within the neuro-physiology of the organism. However, the fact that the organism *senses* the world as “really” existing outside himself needs to be accounted for.

J. RETROJECTION: THE ACT OF “CASTING BACK”

Retrojection (from *retroject*: “To cast or throw back”) is a word so rarely used it might almost be considered a neologism; however, it does describe the act of reference by which the organism automatically (and perhaps only figuratively) casts its sensational experience of the world back upon the world of actual objects and events.

As an act of reference, retrojection does not merely cast back upon the sensated surround the contents of an immediate experience of objects and events, as a mirror reflects only that which is immediately before it. To the indifferent manifold of energies and other properties of the surround which have become translated into electro-chemical excitations of the organism’s receptor-neural means of access, the central-nervous-system process of retrojection contributes certain components from the previous experiences of the organism. Thus the surround of immediate experience is furnished with animation—that “living quality” which the individual human appreciates according to his experiential capacity. To the things-merely-seen are added characteristics of solidity or softness, roughness or smoothness, warmth or coldness, etc., as well as a potential for the induction of pain or pleasure; such additions are contributed from the fund of the organism’s past experience of objects and events—its collisions, manipulations, and other previous transactions with its environment. By means of such experiences of response—with their inevitable accompaniment of sensations—the organism comes to *sense* the externality of its surround. Thus, the world which might seem no more than a solipsistic illusion were the organism to locate it where it is actually sensated comes to be *realized* according to the mandate of response experiences.

It is within the broad conceptual framework of retrojection that the selective function of *stimulus-ascription* operates. The varieties of experience

which the organism continues to survive provide it with a reservoir of residual values which can be more or less faithfully applied to the discriminated aspects of its current events. To the human whose reservoir is full as a consequence of a well-rounded interest in the external world, any surround he chances to discover himself within becomes freighted with potentiality for additional varieties of experience. At the other extreme, to the human whose reservoir remains practically empty by reason of inertia, that which he casts back on the surround may be only a reflection of his moribund state—hardly more able to solicit his responses to the world than is the indifferent manifold of physical energies of the world itself.

When the animate properties of experience are restored to the organism, the semblance of reality remarked in the imaginings and dreams of the respondent—which events take place under circumstances of largely internal causation—is more readily accounted for; if reality *as experienced* continues to reside within the organism, it contains all of the lively properties to be found not only in the usual material of dreams and imagination but also in illusion and delusion. All that is lacking is the external raw structure of the real world which furnishes the formal support for the ordinary waking experience of the respondent, much as the loom confines the tapestry without determining the intricate designs of the weaver.

Since the internal stimulus includes all of the potentialities for the instigation of behavioral responses as a centrally represented function of the organism, the necessity for a circular definition of the stimulation process is overcome. The "external stimulus" is not merely defined in terms of response, it can only be recognized as functional upon the occurrence of response, and is thus a product of an illogical "assertion of the antecedent." The internal stimulus function can be defined in terms of the ascription of stimulus-values, whose separate specifications can be set forth without necessary reference to responses; all such ascriptions have a potential status rather than the inevitable "actual" status of the input-output definition.

Quite obviously, the relocation of the sources of behavioral stimulation within the organism increases the already heavy burden of the behavioral scientist. The relatively simple and quasi-observable external "stimulus" or "stimulus situation" as an instigatory input is banished. In its place, a covert and complex stimulus source is proposed. Yet, if the new location is more obscure, it promises a more fruitful consequence than does a pursuit of mysteries in a place whose best recommendation is adequate illumination.

K. SUMMARY

It has been proposed that the present conceptualization of the *external stimulus* represents an animistic residue within the behavioral sciences themselves—which residue enforces a schematic ordering of the organization of behavior on an “input-output” basis, with responses being instigated by the energies and other properties of the external universe. This assignment of function violates certain long-established principles of physical science, which deny either intrinsic motivation or other forms of animation to the properties of the physical universe. The concept of external stimulation also violates certain basic premises of the biological, philosophical, and psychological sciences—according to which the organism is specifically limited by the nature of its mode of contact with the objects and events of the external universe.

Several observations were ventured to account for the firm entrenchment of the stimulus concept and its associated implications:

1. The misapplication of causal relations to data of the behavioral sciences: if an observer takes notice of a response *qua behavior*, he can legitimately regard that response as an effect; however, should he locate an external *cause* by which he can account for the behavioral effect, he may not lawfully refer to that cause as a “stimulus” for the observed response. Nevertheless, this equation appears to be regularly applied to psychological data within an “input-output” schema.

2. Unwarranted generalization by means of which the term “stimulus” as it has long been applied to a laboratory *instrument* is extended to include certain instigatory powers supposedly resident in the external environment. In its original significance, the term “stimulus” bore no necessary relationship to the “responses” *qua sensations-in-consciousness* for the introspectionist; his subjects were trained to report their conscious sensations. Removed from the original context of “stimulus” (as laboratory instrument) and “response” (as sensations-in-consciousness), the term *stimulus* developed an instigatory rôle having a necessary function in the psychologist’s theoretical accounts of behavior.

3. The misapplication and extension of the terminology and operations of the physiologist’s laboratory: The terms “stimulus” and “response” as used by the physiologist may legitimately apply to his field of operations, but when they are extended into the field of strictly psychological operations, such terms retain no semantic validity—although they may maintain an appearance of validity sufficient to perpetuate their usage as it has become misapplied.

4. The status of the behavioral scientist as an organism among organisms: since he bears a subjective relationship to the material under his study, the behavioral scientist is far more liable to suffer interference with his efforts toward the achievement of objectivity than are such scientists as treat of the purely physical universe.

5. The ahistorical conceptualization of the organism: the static mode of conceptualization through which the organism-in-its-environment may be observed to respond "to" the objects and events of the external surround represents a beguiling *appearance-of-things* which renders plausible all of the foregoing misapplications of the stimulus concept. The ahistorically conceptualized organism seems to respond "naturally" to the animistically defined eliciting properties of the physical universe.

A series of observations were made relative to the actual limitations of the organism as a respondent within a physical surround having the specifications laid down by the physical sciences:

1. The physical universe is an indifferent manifold of energies and other properties which has no power to "elicit," "provoke" or otherwise control the behavioral responses of the organism, and has no discoverable vested interest in its activities such as would induce it to guide or misguide the organism's behavior.

2. It is necessary to establish a clear-cut distinction between statements whose referent is the problem of *access* and statements whose referent is the problem of *significance*; the receptor-means by which the organism is enabled to make contact with its surround are correlative to the energetic (and other) properties of the physical universe, and if the external energies are capable of exciting receptor-neural activity, they thus serve as the instrumental means by which the organism obtains some degree of *access* to the universe. There is no warrant for the assumption that such "stimulation" either furnishes the organism with some form of knowledge of the *significance* of the objects and events of the environment or serves to instigate the behavioral responses of the organism.

Problems of *access* and *significance* relate to separate and distinctly different levels of discourse and when they are unwittingly confounded, the resultant "synthesis" comprises a modern version of the ancient ideolon theory of behavioral stimulation.

3. The energies and other properties of the physical universe which render certain aspects of the environment accessible to the visual, etc., receptor-systems of the organism do not in any manner guarantee the validity of such objects and events as they seem to represent. Although the organism

may act as if certain qualitatively appreciated properties of objects, etc., adhere faithfully to those external objects (as may be inferred from its responses within controlled situations) the scientific observer must be wary of becoming either implicitly or explicitly seduced into sharing its faith in the *appearance-of-things*. If external energies could in any manner guarantee the validity of their objects, such a state of affairs would represent an ideolon-in-fact.

4. Since the energies and other properties of the surround are indifferent to the welfare of the animal kingdom and since the organism cannot infer the significance of such objects and events as bear on its continued well-being from the nature of the inputs of energy which excite its receptor-means of access, it was proposed that the conceptualized sources of behavioral stimulation be relocated within the organism—including such sources as have been designated as “potential stimuli.”

A conceptual framework was finally proposed which would at once account for the appearance-of-things relative to external stimulation and at the same time resolve the problem of effective contact between organism and environment. The schematic order of behavioral organization was hypothetically reversed in terms of a proposition which states that behavioral organization proceeds *from* the organism *toward* the environment and not the other way around as has been traditionally inferred. Further, it was proposed:

1. That the organism operatively (as distinguished from *volitively*) establishes the significance-value of discrete aspects of its surround as a result of its collisions, manipulations and other transactions with the objects and events of the external environment and that in consequence of its various experiences the organism is enabled to ascribe a hierarchy of stimulus-values to such objects and events as somehow bear on its prospects for continued well-being. While the responses of the organism may be addressed toward the environment, its behaviors are stimulated not *by* any of the properties of the external surround, but by the *stimulus-values* which the organism has *ascribed* to the myriad aspects of its world.

2. That the ascription of stimulus-values occurs within the conceptual framework of a process designated as *retrojection* (from the verb, *retroject*: “to cast or throw back”). Since the receptor-neural systems transmit their electro-chemical impulses only centripetally toward a terminus in the central nervous system, the environment *as it is sensed* has its location within the organism. The organism automatically refers the surround-as-sensed back “outside” where its experiences of response indicate it really belongs.

However, by its retrojections, the organism does not merely cast back what it receives, after the fashion of a mirror. To its immediate sensory experience must be added the animate qualities of past experiences which somehow bear relationship to the current event of sensation. The organism thus responds to its own animations of the environment which represent its stimulus-values cast back upon the indifferent manifold of energies and other properties of the external universe.

REFERENCES

1. ALLPORT, F. H. *Theories of Perception and the Concept of Structure*. New York: Wiley, 1955.
2. BORING, E. G. *A History of Experimental Psychology*. New York: Appleton-Century-Crofts, 1950.
3. HULL, C. L. *Principles of Behavior*. New York: Appleton-Century-Crofts, 1943.
4. JAMES, W. *The Principles of Psychology*. New York: Holt, 1890. (Single volume edition: 1950.)
5. LEWIN, K. *A Dynamic Theory of Personality*. New York: McGraw-Hill, 1935.
6. NISSEN, H. W. Phylogenetic comparison. In *Handbook of Experimental Psychology*. (Stevens, S. S., Ed.) New York: Wiley, 1951.
7. SHEFFIELD, F. D., & ROBY, T. B. Reward value of a non-nutritional sweet taste. *J. Comp. & Physiol. Psychol.*, 1950, **43**, 471-81.
8. SHEFFIELD, F. D., ROBY, T. B., & CAMPBELL, H. A. Drive reduction versus consummatory behavior as determinants of behavior. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 349-354.

General Delivery

West Side Station

Colorado Springs, Colorado

THE RELATIONSHIP OF TWO MEASURES OF AUTHORITARIANISM TO THE MINNESOTA TEACHER ATTITUDE INVENTORY*

College of Education, University of Georgia

ALBERT J. KINGSTON AND GEORGE L. NEWSOME

A. PROBLEM

The Minnesota Teacher Attitude Inventory (2) has been widely used to predict teaching efficiency and aptitude (3, 4). The inventory was developed empirically and was designed to determine how well teachers get along with their pupils by measuring the attitudes which teachers are likely to bring to their classrooms. Although the inventory contains items relating to a wide variety of attitudes, many items appear to measure attitudes which reflect on the interpersonal-relationships between pupil and teacher. Such items may be regarded as measuring opinions and attitudes which are reflective of authoritarian patterns of behavior or belief. The present study is designed to illustrate the degree of relationship between teacher attitudes as measured by the MTAI and opinions and attitudes which reflect authoritarian opinions and attitudes as measured by two independently developed instruments.

The WSF Scale (8) consists of items grouped into clusters according to the theoretical design employed by Adorno, Frankel-Brunswick, *et al.* (1). The original items were taken from various personality tests and all of the items correlated significantly with the F Scale employed by the authors of *The Authoritarian Personality* (1). The authors of the WSF Scale believe the instrument to be more personality centered and less ideological than is the F Scale. The scale itself consists of 149 items and the subject is merely asked to indicate whether each statement is true or false. Unlike the other instruments employed in this study, a high score indicates a tendency toward authoritarianism.

The Inventory of Classroom Administrative Philosophy (5) was designed to determine the degree to which teacher's practices and beliefs reflect the concept of democracy. It consists of 150 items evolved from democratic principles. Items include a wide range of classroom practices as well as

* Received in the Editorial Office on February 23, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

beliefs concerning children's behavior and attitudes. The subject is asked to agree or disagree with each statement.

B. PROCEDURE

A group of 79 female white elementary teachers who were enrolled in a special in-service training workshop were employed in the study. Each of the teachers was administered all three instruments. The group ranged in age from 21 to 59 years of age and from one to 34 years in teaching experience. They represented 10 different school systems. The group was felt to be representative of the county school systems in the state from the viewpoint of age, experience, and certification.

C. RESULTS

The mean scores on each instrument indicated that as a group, the teachers tended to be somewhat conservative and domineering. The median score attained by the group on the MTAI was below that of the norms furnished by the test authors. Although comparable groups were not available for a comparison of scores on the WSF and ICAP, the scores on these instruments were found to be somewhat less than expected. Table 1 illustrates the mean scores and standard deviations attained on each instrument.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF EACH SCALE
(N = 79)

Instruments	Mean	σ
MTAI	20.60	30.36
ICAP	104.10	12.70
WSF	82.87	25.50

As shown in Table 1, the ICAP scores revealed a greater homogeneity than did either the MTAI or the WSF. These results may be due to the wording of the items which tended to be couched in educational terms which may have enabled the teachers to "see through" the items. On the other hand, the WSF Scale posed a threat to certain teachers. In the discussion held by the authors after the tests were administered, a number of the participating teachers mentioned that they felt some of the WSF items asked questions which were no one's business. The majority, however, felt that the ICAP was a fair and seemingly valid measure.

Intercorrelations among the three instruments were obtained and all coefficients were significant at the one per cent level of confidence. Table 2 illustrates the obtained coefficients.

TABLE 2
INTERCORRELATIONS AMONG INSTRUMENTS
(N = 79)

Instruments	γ	ρ
MTAI - ICAP	.312	.01
MTAI - WSF	— .428*	.01
ICAP - WSF	— .480*	.01

* A negative correlation is due to the method of scoring used with the WSF Scale. A high WSF score indicates authoritarian personality while on MTAI and ICAP a high score is indicative of a democratic attitude.

In an effort to determine the nature of the items which discriminated, an item analysis was undertaken, employing the method described by Stanley (6, 7). The highest and lowest 21 individuals were utilized in this procedure. Sixty-four items from the WSF, 63 from the MTAI and 39 from the ICAP yielded values at the 2.5 per cent level of significance. These results tended to point out the greater discriminating ability of the WSF and MTAI over the ICAP previously noted. Table 3 lists the items from each scale which discriminated at the 2.5 per cent level.

The aid of several colleagues was solicited to form a jury to place the discriminating items from the MTAI and ICAP into the categories employed in developing the WSF. As little agreement among the jurors was obtained, this approach was abandoned. A careful examination of the discriminating items, however, revealed certain similarities and overlapping, and it was felt that many of the items were measuring the same basic personality factor.

The scores were then analyzed to determine the effects of age, experience, and formal education upon the responses of the group. No significant mean differences were found.

D. SUMMARY AND CONCLUSIONS

A comparison of the MTAI with two scales which presumed to measure authoritarianism yielded significantly positive correlations. Teachers who scored high on the MTAI tended to score high on the Inventory of Classroom Administrative Philosophy. Teachers who scored high on the MTAI tended to score low on the WSF Scale indicating a tendency toward democratic beliefs. As the WSF Scale represents a personality-centered instrument, it may be concluded that scores on MTAI tend to reflect certain personality characteristics indicative of authoritarian behavioral patterns. The significant relationships found among the instruments may be indicative that all three instruments tend to appraise some aspects of an authoritarian "syndrome."

An examination of the items which discriminated on each of the instru-

TABLE 3
ITEMS WHICH DISCRIMINATE

Item No.	Item No.	Item No.	Item No.	Item No.	Item No.
Minnesota Teacher Attitude Inventory*					
2	38	52	80	106	133
7	39	57	82	109	137
12	40	58	85	110	139
13	42	60	86	111	143
14	43	61	87	113	144
15	44	65	90	116	145
17	46	68	93	119	149
19	48	70	95	122	150
21	49	74	98	126	
31	50	75	104	128	
36	51	76	105	132	
WSF Scale*					
6	35	52	82	112	131
8	37	53	86	116	132
10	40	55	89	118	133
16	42	58	95	120	139
18	43	61	99	121	140
26	45	62	102	122	141
27	46	63	103	123	144
28	47	66	107	124	145
32	48	67	109	125	147
33	50	68	110	127	
34	51	69	111	130	

Inventory of Classroom Administrative Philosophy

1. I think that the role of the teacher should be to organize the materials of instruction but not to participate actively in their development.
13. I believe that the giving of additional study assignments as punishment has a negative effect upon the student.
14. I think that most adolescent children will lie or tell an untruth in order to avoid unpleasant consequences due to their misconduct.
17. I believe that a student should be discouraged from constantly seeking approval from the teacher rather than from his classmates.
21. I feel that students should not question the grades which a teacher gives them.
26. I feel that teaching would be much more effective and satisfying if you could expel certain students just as they discharge unsuited employees in industry.
29. I believe that students with low IQ's should be taught (in a nice way) that they cannot expect to get as much out of life as a brighter student can expect.
33. I think that adolescent children would rather be praised by the teacher than by their fellow students.
35. I believe that a self-government program in the classroom would be too time-consuming to be practical in the school.
37. I think that one of the best ways to discipline an adolescent child is to inform his parents of his misconduct.
41. I think that students who are continually in trouble should be expelled for the good of other students.
46. I think that a rigid classroom schedule helps to discipline a student's learning.

* These items are listed by number because they are fairly well known instruments. See references for further information.

TABLE 3 (Continued)

Inventory of Classroom Administrative Philosophy

47. Because students are apt to be restless if not occupied, I think that the teacher should even go so far as to provide work for the purpose of keeping them busy.
53. I think that a teacher should not answer, "I don't know" to a student's question.
56. I think many students deliberately try to misinterpret instructions.
62. I believe that a teacher should become friendly with his students, in spite of the fact that they may take advantage of him.
64. I think that one fault of student self-government programs is that students tend to act old too soon.
68. I think that a child usually does what he thinks is right in a situation.
69. I believe that I did not constantly assert my authority, the students of the class would tend to get out of hand.
72. I think that students often act like "little savages."
79. I think that most of the infractions of school rules made by youngsters result from their misunderstanding of the purpose of the regulations.
85. I think that most pupils who are talkative should be discouraged.
86. In my opinion, a good way to maintain discipline is to keep the trouble-makers isolated from the rest of the class.
87. I believe that all students should be allowed major management of their extra-curricular activities and school clubs.
88. I prefer to teach students whose parents are not foreigners.
96. I think that it is wise to encourage the pupils of the class to inform me of any misconduct on the part of other members of the class.
98. I believe that if they keep forcing us to teach everyone's children in our schools, a large group of poorly-educated people will eventually result.
100. I believe that once the teacher makes up a work schedule, she should follow it closely.
103. I think that the fear of being caught should be employed to make students obey the school regulations.
107. I am afraid that some children need some type of discipline which is more stringent than the present school law allows, in order to make them behave.
114. I feel that controversial issues should not be taught in school.
115. By sending a student who continually disrupts the classroom activities to the principal, I believe that the teacher is admitting her inability to handle the classroom activities.
122. I believe that enforcing discipline is one of the most time-consuming duties of the teacher.
123. I think that if students cannot appreciate the things that an education offers, they should be allowed to leave school, regardless of their age.
128. I am certain that the children of parents of certain nationalities are usually less intelligent than other children.
129. I try to take an objective, impersonal attitude toward students who violate classroom regulations.
130. I think that the trouble with much of the group activities that we have tried in the schools, is that children often enjoy flouting the proper school authorities.
138. I believe that students should be given the opportunity to evaluate their own work at the time grades are assigned.
146. I have observed that my pupils tend to be rather apathetic about tackling new work topics.

ments is included for further study. The similarity and overlapping of content among the discriminating items reveals that all three instruments in part may be measuring similar or identical patterns of personality.

REFERENCES

1. ADORNO, T. W., FRANKEL-BRUNSWICK, E., LEVINSON, D. V., & SANFORD, R. N. *The Authoritarian Personality*. New York: Harper, 1950.
2. COOK, W. W., LEEDS, C. H., CALLIS, R. *Minnesota Teacher Attitude Inventory*. New York: Psychological Corp., 1951.
3. DAY, H. P. A study of the predictive validity of the Minnesota Teacher Attitude Inventory. *J. Ed. Res.*, 1959, **23**, 37-38.
4. HOYT, C. J., & COOK, W. W. Predictive validity of the Minnesota Teacher Attitude Inventory based on pupil attitudes toward school. *J. Teach. Educ.*, 1959, **10**, 42-44.
5. KINGSTON, A. J. *An Exploratory Study of the Concept of Democracy as Reflected by the Teaching Practices and Beliefs of Secondary School Teachers*. (Unpublished Doctoral Dissertation, Cornell University, 1950.)
6. ROSS, C. C., & STANLEY, J. C. *Measurement in Today's Schools*. (Revised Edition.) New York: Prentice-Hall, 1954. (Pp. 436-453.)
7. STANLEY, J. C. A simplified item-analysis procedure. *Amer. Psychol.*, 1951, **6**, 369.
8. WEBSTER, H., SANFORD, R. N., & FREEDMAN, M. A new instrument for studying authoritarianism in personality. *J. of Psychol.*, 1955, **40**, 73-84.

College of Education
University of Georgia
Athens, Georgia



BEHAVIORAL RESEARCH AND INTERGROUP COMMUNICATION*

Department of Psychology, University of Arizona

WILLIAM J. MACKINNON

A. INTRODUCTION

Behavioral scientists have created a variety of research methods and have discovered a number of empirical principles which bear upon human communication. The actual control of mass communication in America, however, is often one-sided and manipulatory (6, 13). Faced with growing misuses of power, the behavioral scientist may nevertheless conceive of constructive alternatives (14).

One purpose of this paper is to imagine some systems of communication which have become possible largely by procedures developed in the behavioral sciences. Another is to estimate the probable contributions of these systems to goals such as free expression and an interest in the communicative process. The discussion will lead to several suggestions for future research.

B. SOME POSSIBLE QUALITIES OF INTERGROUP COMMUNICATION

In this section, certain characteristics which intergroup communication might have will be considered. The characteristics are somewhat ideal and qualitative. In later sections two particular forms of intergroup communication will be described. These will include some of the more essential characteristics of the present model but will be more precise in meaning and more economical to operate.

Either the ideal or the two "practical" systems of intergroup communication might operate in the local community, the large organization, the nation, or even in the world community. The individuals involved in the communication are assumed to understand the general nature of the process.

The first of the three stages of this process involves individual senders of messages and the individual messages they communicate. The second is the formation of some kind of "group message" from the individual messages. The third focuses on these group messages and the individual receivers who interpret them. The entire three-stage process then starts anew and continues to operate in cyclical fashion.

* Received in the Editorial Office on February 23, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

The general characteristics of the ideal model will be described in terms of these three stages. As the description proceeds, it will become apparent that intergroup communication with some of these characteristics is already appearing. It occurs to some extent in social-psychological surveys (4) and clinical-group-dynamics studies (1), in research which combines both survey and clinical aspects (9, 10), and in the reactions to such investigations. The following then are the general characteristics of an ideal intergroup communication process.

1. *Individual Senders and Individual Messages*

The individual members of groups send their messages under assurance of personal anonymity. Whether writing messages or dictating them, whether making statements or asking questions, the communicators do not reveal their personal identity, nor do they refer to specifically identifiable individuals in the communication system. They address their message to a group rather than to named persons. They identify themselves or permit themselves to be identified only as members of a group. The groups, whether senders or receivers, are either demographic, e.g., manual workers and non-manual workers, or socio-psychological, e.g., working class and middle class as defined by verbally expressed feelings of belonging. A group might consist of all members of the total communication system or only of a subgroup. Senders request to see any group messages which interest them, including messages not directed to the groups to which they belong.

2. *Individual Messages and Group Messages*

Group messages replace individual messages as the content actually received by individual interpreters. These group messages are as representative as is feasible of the total individual messages obtained from the group members. They would not be very representative unless all members of the group had produced messages, because communicating changes the beliefs and attitudes of the communicators. For any group, representativeness or a good approximation to it is secured by random sampling from all individual messages sent by its members, or by a content analysis of this population of messages, or by a combination of random sampling and content analysis. The group message includes the context provided by the group identifications of the senders and receivers.

3. *Group Messages and Individual Receivers*

Three factors determine the particular individuals who receive particular group messages. One is the senders' choices of target groups. These choices

might be influenced by the second factor, questions put to the senders in previous messages. The third factor is the desire of the individual receiver to study the messages sent by any group to any destination. Because group messages are representative of individual messages, the views of different groups can be compared. As was said earlier, the intergroup communication process is circular. Receivers feed messages back to senders, and so on.

C. A DISCUSSION OF TERMS AND CONCEPTS

To prepare the way for a discussion of two more nearly practicable models or methods of intergroup communication, the meanings of a few terms will be considered. Groups, topics, messages, media, situations, and research will be discussed in order.

1. *Groups*

Groups may be unambiguously and publicly defined as in terms of a standardized system of job classifications. On the other hand they may be ambiguously or idiosyncratically defined. When someone speaks of "working people," for example, it is unclear whether or not he intends to include white-collar workers in his concept.

Related to this distinction is the number of groups involved in communication. The set of groups originating or receiving group messages could be fixed at some relatively small number. It could also vary widely, as when it includes any group which an individual could conceive.

An important distinction is one between interest groups and reference groups. The former have some common interest or goal and a consciousness of belonging, which varies in intensity among different members (5). Groups of the second type are referred to when a person seeks a norm or standard to guide his own decisions as to what is true or desirable.

2. *Topics*

A communicator generally has in mind some topic of communication. The topic may be one that is unambiguously and publicly defined, for example, whether a particular legislative bill should be enacted. In contrast, the topic might have somewhat ambiguous or idiosyncratic aspects, as when a communicator refers to "the welfare of old people." Ambiguity is in fact a matter of degree, as are several other distinctions being drawn. All communications content may be at least somewhat ambiguous, or subject to different interpretations (8, p. 83).

Again, the topic may be fixed or it may be allowed to vary, as when each communicator chooses his own subject matter.

3. *Messages and Media*

For simplicity's sake messages are conceived herein as being written or printed; thus they might be read by or to their receivers.

A useful distinction is that between minimal and extended messages. A minimal message is an elementary unit. It may be a single question or statement, often containing less than 10 words. In one study a meaning-unit was defined as "the smallest piece of content that makes an assertion." The sentence, "The intelligent Roosevelt will keep us out of war," was regarded as containing two meaning-units, namely, "Roosevelt is intelligent" and "Roosevelt will keep us out of war" (3, p. 138). The minimal message also bears a relationship to what has been called "a single communicative act." This is often a simple sentence, though sometimes either the subject or predicate will be merely implied. In cases of double predication, the sentence contains more than one communicative act (2, p. 37). The minimal message is obviously atomistic.

In contrast, an extended message is a composition between 100 and 200 words in length, an essay bearing upon a single topic. It has a configurational property; its parts, fundamentally minimal messages, cohere meaningfully.

Whether group messages consist of detached minimal messages or the more articulated extended messages, they include mention of the sending and receiving group as context. The interpreters of group messages receive some indication of the identity of the sending group.

Messages might be transmitted over a common medium or only over a special medium. In other words, all members of the communication system might receive a particular group message, whether or not it was addressed by the senders to all members, or else only a subgroup to whom the message was addressed might receive it. Relevant in this connection is the choice of translating group messages into a common basic vocabulary or leaving them more or less as originally formulated in individual messages.

4. *Situations*

Intergroup communication could occur in at least two different general situations. In the first the group is about to make a decision regarding the distribution of valued objects among the members, and it will be able to control the distribution with social, economic, or physical power. Such a state of affairs will be called a group-value situation. It contrasts with individual-value situations. Here individuals are approaching a choice between personal values, as for example, between different paths or styles of life; and, hypothetically, public opinion or reaction does not favor one alternative over

another. Few or many persons may be involved in making such choices at any given time.

5. *Research*

In connection with intergroup communication, research could be of two types. Realistic research would study an actual system of intergroup communication while it was in progress. Artificial research would pertain to individuals requested to act *as if* they were participating in a real system of communication which might be difficult to bring about, for example, to compose messages upon the false assumption that they might be published in a world-wide medium.

D. UNSTRUCTURED ANONYMOUS INTERGROUP COMMUNICATION AND RELATED RESEARCH PROBLEMS

The first of the two practical models to be described involves some highly unstructured elements. Individual senders compose their own extended messages, select the target groups in their own terms, describe as they wish the groups to which they belong, select any topic on which to communicate, and define it as they wish. In all likelihood the groups and topics would be multifarious and the definitions frequently ambiguous.

A large random sample of all the extended messages thus composed is reproduced in a common medium in a common basic vocabulary. Further stockpiles of random samples of untranslated extended messages are available for examination by any member of the communication system. Only after extensive perusal of many such messages, including the sender-receiver contexts, would an individual be likely to discriminate representative differences between the topics broached by different groups, or between the views held on a single topic by various publics. In so far as the participants can sense such collective differences, the system is one of group communication.

Research might decide this question, as it might provide answers to many others. What topics would be more prominent in which groups? Where would the topic of international relations stand? What drift in topic frequency would occur over time? Would individuals say what they really wanted to, barring, of course, communication content proscribed by law? What would laboratory apparatus indicate as to truthful report if subjects were asked, "Apart from illegal communications, did you write (or say) what you really wanted to?" or "Did you feel that your anonymity was assured?" How would these factors change with repeated sequences of sending and receiving messages?

E. STRUCTURED INTERGROUP COMMUNICATION AND RELATED RESEARCH PROBLEMS

Another possibility for intergroup communication presents a high degree of structure. The communicators confine their comments and queries to a fixed, unambiguous, public topic. The principal interest groups or reference groups are likewise fixed, unambiguous, and public. Thus the topic might be "the forthcoming contract between union and management in industry *X*," and the principal groups might be the union members, the management, and the general public.

If only two main groups, *A* and *B*, are communicating, messages could travel in four directions: from *A* to *B*, from *B* to *A*, from *A* to *A*, and from *B* to *B*. (Intragroup communication, as from *A* to *A*, is often intimately related to intergroup communication, as when members of *A* decide what messages should be sent from *A* to *B*.) In the instance of two groups, therefore, there may be four group messages at each stage of senders-to-receivers cycle of the communication process. If the number of groups increases to three, the number of directions or routes that group messages could take increases to nine. There may now be nine group messages at each cycle.

To prepare the group message, a random sample of extended messages sent by all members of the group is selected. The extended individual messages from, say, members of *A* to members of *B*, are next divided into minimal messages which are then sorted into classes, all units having the same meaning going into the same category, and thereafter being represented by a standard wording. The units in each class are counted and a frequency distribution arranged. Since the individual communicators composed extended messages, the interviewer-bias effects of public opinion research would be absent. The test for reliable categorizing of content is agreement among judges, some of whom share and some of whom disown the views expressed in the content. There is, of course, a "miscellaneous" or catch-all category for those minimal messages with very low relative frequencies. Samples are sufficiently large so that reported percentages would very probably be highly representative of corresponding percentages in the populations. Finally, the group messages are presented to all members of the communication system in a common medium and basic vocabulary.

A large number of additional permanent interest or reference groups communicate. These groups are fixed, unambiguous, public, and somewhat overlapping. They produce messages of a special sort merely by checking preferred minimal messages which represent their viewpoints. The minimal messages from which they choose are those obtained in the debate between

the principal groups except that they do not contain their original sender-receiver context. The list of message units from which to choose would probably contain statements meaningfully grouped as, "The worker's wages are too low"; "the worker's wages are too high"; "the investor's profits are too low"; "the investor's profits are too high"; "a fact-finding committee should be appointed"; "a fact-finding committee should not be appointed." Any individual would receive the opinions, suggestions, etc., of any standing groups whose views were of interest to him. Graphic presentation of the relative frequencies would expedite comparisons of the various group standpoints.

Many questions for research arise in connection with this method or model for intergroup communication. How reliably could content be unitized and categorized? From a predetermined list of fixed, unambiguous, public reference groups, which ones would be selected by subjects most often as the groups whose messages they would hypothetically wish to examine? What would be the relative standing of clergymen, ethicists, psychiatrists, psychologists, social-welfare workers, sociologists, economists, political scientists, artists, novelists, manual laborers, white-collar workers, small-business owners and operators, large-business owners and executives, assuming these various groups were defined in relatively objective terms? How would this hierarchy of interest groups or reference groups vary with the group of choosers and with the topic of communication? What groups would loom prominently as reference groups if the topic of communication were to find an answer to the question, "What do you think is the best thing that the United States could do to help keep peace in the world?" (7, p. 121). In this artificial or as-if type of research, would choices differ depending upon whether they were obtained under classroom or clinical conditions?

F. PROBABLE EFFECTS OF THE HYPOTHETICAL COMMUNICATION SYSTEMS

The communication possibilities which have been suggested, whether ideal or practical, would probably have psychological effects. Behavioral research of the past suggests two such results, an increase in free expression and a growth of communication interest.

1. *Free or Voluntary Expression*

The condition of anonymity stands in the foreground as a method of reducing counterforces or barriers to expression. That such counterforces exist is well demonstrated by Stouffer's research on civil liberties (16).

Majorities or near majorities of the American population would not allow a person to speak in the community if he took a certain viewpoint on some topics. Yet it appears from Center's earlier survey that some Americans do hold at least some of these viewpoints which many would proscribe (4). Now the reduction of opposing forces sometimes leads to desired behavior with less resultant tension than does an increase in the driving forces themselves (12).

The immediate effect of promising a person anonymity might partially remove the threat of economic and social penalties and thus lead to freer expression. However, the communicator might still restrain some views which he felt were unique to himself, and contrary to the universal or nearly universal views of others. If he and others like him found these views expressed to some extent in the anonymous communications they received, the state of "pluralistic ignorance" would dissolve, and the inhibited communicators might thereafter express their private attitudes in their further communications. A comparable prediction was made in an early study of private and public attitudes (15, p. 101).

2. *Tolerant or Interested Participation*

If interest in political or ethical communications in the mass media is not keen at present, it may be because the content is one-sided. Those whose interests are favored scarcely need take account of opposing arguments which never or rarely are communicated in the mass media. In turn those whose expression is thus blocked can only despair.

If, in social-value situations, political topics were debated as in the models which have been discussed, a new interest might arise. The discontented and disadvantaged would have an outlet for expression. Their hitherto aloof antagonists would probably rally to their own verbal defense. The debate should be as vitally interesting as the subtitle to a well-known book on politics, namely, *Who Gets What, When, How* (11).

Even when, as in individual-value situations, where no legal force or public pressure is involved, interest in anonymous communication could run high. People concern themselves with the individual choices made by others, since the values and cognitions involved in these choices are relevant to their own decisions.

Another, currently common, situation is that in which group norms are changing, and private practices or standards may not align with them. Here questions of risk and conflict arise. Thus anonymous intergroup communi-

cation would probably contain some of the vital and intriguing qualities of discussions in group psychotherapy.

Communication, furthermore, often generates group organization among persons of similar instrumental or consummatory interests. The groups in turn mobilize impulses to action which could be expressed in anonymous, intergroup communication.

G. SUMMARY AND CONCLUSION

Communication systems of various sorts have been described. They have as common characteristics the personal anonymity of the individual communicators and intergroup messages which are largely representative of views in the participating groups. The provision of anonymity might do for communication what the secret ballot has done for voting—free it from the restraining forces of social-psychological and economic sanctions. Focussing attention upon liberal possibilities in intergroup communication suggests research which itself might improve the chances of the methods being applied.

REFERENCES

1. BACH, G. R. *Intensive Group Psychotherapy*. New York: Ronald, 1954.
2. BALES, R. F. *Interaction Process Analysis*. Cambridge, Mass.: Addison-Wesley, 1950.
3. BERELSON, B. *Content Analysis in Communication Research*. Glencoe, Ill.: Free Press, 1952.
4. CENTERS, R. *The Psychology of Social Classes*. Princeton: Princeton Univ. Press, 1949.
5. ———. The intensity dimension of class consciousness and some social and psychological correlates. *J. Soc. Psychol.*, 1956, **44**, 101-114.
6. Commission on Freedom of the Press. *A Free and Responsible Press*. Chicago: Univ. Chicago Press, 1947.
7. COTTRELL, L. S., JR., & EBERHART, S. *American Opinion on World Affairs in the Atomic Age*. Princeton: Princeton Univ. Press, 1948.
8. FEARING, F. Toward a psychological theory of human communication. *J. Pers.*, 1953, **22**, 71-88.
9. KINSEY, A. C., POMEROY, W. B., & MARTIN, C. E. *Sexual Behavior in the Human Male*. Philadelphia: Saunders, 1948.
10. KINSEY, A. C., POMEROY, W. B., MARTIN, C. E., & GEBHARD, P. H. *Sexual Behavior in the Human Female*. Philadelphia: Saunders, 1953.
11. LASSWELL, H. D. *Politics: Who Gets What, When, How*. In the *Political Writings of Harold D. Lasswell*. Glencoe, Ill.: Free Press, 1951.
12. LEWIN, K. *Field Theory in Social Science*. (Ed. D. Cartwright.) New York: Harper, 1951.
13. PACKARD, V. *The Hidden Persuaders*. New York: McKay, 1957.
14. ROGERS, C. R., & SKINNER, B. F. Some issues concerning the control of human behavior. *Science*, 1956, **124**, 1057-1066.

15. SCHANCK, R. L. A study of a community and its groups and institutions conceived of as behaviors of individuals. *Psychol. Monog.*, 1932, **43**, No. 2 (Whole No. 195).
16. STOFFER, S. A. Communism, Conformity, and Civil Liberties. Garden City, New York: Doubleday, 1955.

Department of Psychology

University of Arizona

Tucson, Arizona

THE STUDENT'S NEED FOR AN ISSUE*

Department of Psychology, University of Florida

MALCOLM H. ROBERTSON¹

The story of how people come to write anything is probably as informative (and sometimes more entertaining) than what they write about. I recall very well the incident that prompted me to write this article. I was on my way to lunch when I met a psychologist whose acquaintance I had made only recently. We quickly fell in step and after a few token efforts to avoid shop talk we settled down comfortably to a rather discursive conversation about our respective jobs. It was not long before he inquired about the state of my research interests and I described one or two ideas that I had rehearsed for just such an occasion. I must have sounded apologetic for he suddenly interrupted me to say that he did not really care what a person was interested in just so he was enthusiastic about something—anything!

This remark must have struck a familiar note for almost immediately I thought of a day several years before when, as a graduate student, I was listening to a professor giving advice on how to get started on a thesis. I remembered nothing of what the professor had said with the exception of one rather astonishing remark, "pick something close to your heart." Aside from being appalled by such an unscientific sounding phrase (I had just completed my first year of graduate study and was still quite sensitive to such language), I really had no particular reason to recall this comment.

However, at this point I returned to my friend who by then was several paragraphs ahead of me. Although our discussion soon veered off to tangential topics, I still wondered if he appreciated the impact of his remark and how incisively he had undercut a philosophy of training being held by a large number of psychologists responsible for graduate instruction. In any case, my ruminations did not end there. Instead, I found that with a little more consideration I was able to pull together a number of scattered observations and thoughts which seemed to dovetail with the comment of my companion.

* Received in the Editorial Office on February 25, 1960, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹ I am grateful to Dr. H. L. Ansbacher for his suggestion of the title for this article. However, the opinions expressed in the paper are entirely those of the author.

For instance, out of the misty recollections of graduate school days, a few impressions could be culled. One particularly strong impression had to do with bull sessions among fellow students wherein they would talk about their intellectual rootlessness and nomadism. In an unguarded moment, they would admit to the feeling that they were missing a meaningful experience by not being a strong believer or follower of some school of thought. They would often confide a wish to belong to some group who championed a theory. They longed to become enthusiastic or excited about something, anything, to have a strong identification with some "issue" in psychology. They felt that it was all very fine to be eclectic and sample all the points of view while remaining detached and unimpressed. But, then, there was the "left out feeling." Perhaps there was something of value in being a part of some movement, a defender of some cause, even though it might turn out to be the wrong one. One is reminded of Hall's and Lindzey's (1) plea for the student "to immerse himself in one theory."

Yet I know many psychologists who would label such a course of action as hopelessly out of date, a giant step backwards. They would be quick to invoke the specter of intellectual values being subordinated to rank proselytism. Well, perhaps it is wiser and in the better interests of students not to encourage them to become enamored of one doctrine or to restrict themselves to one point of view.

On the other hand, such admonitions, gentle and subtle though they may be, sometimes overshoot the mark and the student avoids becoming preoccupied with any problem or issue. Maybe there is no harm and even some benefit in a temporary attachment to a particular theoretical orientation. It could even be viewed as a necessary stage in the professional growth with a transition from being devoted to some theory to being devoted to some topic.

As a matter of fact, there is some reason to believe that the capacity to become inspired about some area of investigation must be nurtured during graduate training if it is to become self-perpetuating in the future. In any case, it is at this stage of professional development that one reveals the susceptibility for strong identifications and sudden enthusiasms which later on are often buried under administrative routines and busy work as the person becomes settled and just kind of tired.

The desire to have a stake in some psychological problem or approach is probably a rather common characteristic but the optimum time to spark it is during the latter part of graduate study. However, it is often during this time that many students, if not compelled, are at least counseled to renounce such behavior as unbecoming. Instead, they are advised that to

explore and savor fully while remaining theoretically uncommitted is the hallmark of intellectual maturity. This is the gospel of eclecticism—safe, sane, secure. What budding loyalties or inspirations have been sacrificed to this hallowed term! What injustices have been perpetrated in its name, at least in the unwarranted extension of the term!

To keep the student on the straight and narrow when he is not under the protective custody of his professors, he is given the image of the typical devotee, a person who is singleminded, obsequious, and urgently in need of intellectual emancipation. He is invited to compare this image with that of the eclectic, a person of broad theoretical allegiances, scorning any restrictive ties.

Before long the student is convinced that there is something wrong with wanting to be a part of some new movement or school of thought. Pretty soon he feels guilty about being an ardent supporter of anything. Perhaps he remembers the day when he came across the "penetrating insights" of Freud or experienced a "sudden illumination" upon reading Lewin, or discovered that in Allport's principle of functional autonomy lay the key to the neurotic paradox. He may also recall how he could hardly wait to talk to one of his professors in order to share these revelations. But somehow it didn't turn out the way he planned. Why did the professor have to sound so condescending? Why did he seem so edgy during their talk? Was he just a bit jealous that he was not the one to point out these things? Or, why did he have that triumphant smile? Was it part of that smug feeling about the intellectual vulnerability of students and the fact that such inroads could not be made on his own intellectual composure?

Be that as it may, what is the student, or the newly graduated psychologist for that matter, to gain by developing a strong attachment to some topic or theoretical system? One advantage is that it deepens his professional identity and in doing so it confers a definite direction to his professional development and adds new dimensions of significance to what comprises a very important and a very large part of his life.

One derivative of the deepening professional identity is that he develops a bond not only with those with whom he works but with psychologists in other places and times. His professional affiliation reaches beyond the boundaries of his place of work and study. He knows that there are people who care about what he is doing and what he is thinking. There is a frame of reference for being remembered by people who might forget everything else about him.

Lest the reader get the impression that this is another pitch for togeth-

ness, let me quickly explain that this is definitely not so. For this objective, if pursued vigorously, should arouse enough competition and antagonism to satisfy even the most factious.

Another value, admittedly not a very lofty one, is that often as much recognition can be gained from a few publications dealing with one topic or theoretical framework than from twice as many publications covering a number of areas. To some extent, research becomes easier since each investigation acts as a frame of reference for those to come and one is led quickly (and painlessly) into additional research.

To sum up, manifesting a strong purpose or goal can best be expressed as a sense of dedication which is a mark of professional maturity. The question then is how can this sense of purpose be cultivated in the graduate student or the brand new Ph.D. Probably it is contagious and those who are his teachers and supervisors provide a ready model for imitation. Although he may feel in the beginning that their interests are alien to his, nevertheless he may be impressed by their sense of dedication and be more likely to develop such behavior at a later time.

Moreover, a sense of dedication seems to flourish best in the absence of any discouragement or criticism concerning what inspires or impels the student. More importantly, it is the subtle day-to-day encouragement and support given to him in his intellectual odyssey in which he turns first this way and then that way until at last he finds something that has a special meaning for him and around which his energies can be efficiently organized.

A specific way to implement the above suggestions is through the use of research teams comprised of a faculty member and those students who share the faculty member's objectives. The research teams seem to give a graduate department a more spirited atmosphere through the stimulation that comes from the exchange of ideas and the competitive spirit generated by such groups. Another plan is to have a departmental research project which provides a rallying point for faculty and students and a valuable outlet for rivalries and other interpersonal tensions.

REFERENCE

1. HALL, C., & LINDZEY, G. *Theories of Personality*. New York: Wiley, 1957.

Department of Psychology
University of Florida
Gainesville, Florida





